

A Markov Chain Monte Carlo for Galactic cosmic-ray physics

I. Method and results for the Leaky-Box Model

A. Putze¹, L. Derome¹, D. Maurin², L. Perotto^{1,3}, and R. Taillet^{4,5}

¹ Laboratoire de Physique Subatomique et de Cosmologie LPSC, 53 avenue des Martyrs, Grenoble, 38026, France

² Laboratoire de Physique Nucléaire et des Hautes Energies LPNHE, Tour 33, Jussieu, Paris, 75005, France

³ Laboratoire de l'accélérateur linéaire LAL, Université Paris-Sud 11, Bâtiment 200, B.P. 34, 91898 Orsay Cedex, France

⁴ Laboratoire de Physique Théorique LPTH, Chemin de Bellevue BP 110, 74941 Annecy-le-Vieux, France

⁵ Université de Savoie, Chambéry, 73011, France

Received / Accepted

ABSTRACT

Context. Propagation of charged cosmic-rays in the Galaxy depends on the transport parameters, whose number can be large depending on the propagation model under scrutiny. A standard approach for determining these parameters is a manual scan, leading to an inefficient and incomplete coverage of the parameter space.

Aims. In analyzing the data from forthcoming experiments, a more sophisticated strategy is required. An automated statistical tool is used, which enables a full coverage of the parameter space and provides a sound determination of the transport and source parameters. The uncertainties in these parameters are also derived.

Methods. We implement a Markov Chain Monte Carlo (MCMC), which is well suited to multi-parameter determination. Its specificities (burn-in length, acceptance, and correlation length) are discussed in the context of cosmic-ray physics. Its capabilities and performances are explored in the phenomenologically well-understood Leaky-Box Model.

Results. From a technical point of view, a trial function based on binary-space partitioning is found to be extremely efficient, allowing a simultaneous determination of up to nine parameters, including transport and source parameters, such as slope and abundances. Our best-fit model includes both a low energy cut-off and reacceleration, whose values are consistent with those found in diffusion models. A Kolmogorov spectrum for the diffusion slope ($\delta = 1/3$) is excluded. The marginalised probability-density function for δ and α (the slope of the source spectra) are $\delta \approx 0.55 - 0.60$ and $\alpha \approx 2.14 - 2.17$, depending on the dataset used and the number of free parameters in the fit. All source-spectrum parameters (slope and abundances) are positively correlated among themselves and with the reacceleration strength, but are negatively correlated with the other propagation parameters.

Conclusions. The MCMC is a practical and powerful tool for cosmic-ray physics analyses. It can be used to confirm hypotheses concerning source spectra (e.g., whether $\alpha_i \neq \alpha_j$) and/or determine whether different datasets are compatible. A forthcoming study will extend our analysis to more physical diffusion models.

Key words. Methods: statistical – ISM: cosmic-rays

1. Introduction

One issue of cosmic-ray (CR) physics is the determination of the transport parameters in the Galaxy. This determination is based on the analysis of the secondary-to-primary ratio (e.g., B/C, sub-Fe/Fe), for which the dependence on the source spectra is negligible, and the ratio remains instead mainly sensitive to the propagation processes (e.g., Maurin et al. 2001 and references therein). For almost 20 years, the determination of these parameters relied mostly on the most constraining data, namely the HEAO-3 data, taken in 1979, which covered the $\sim 1 - 35$ GeV/n range (Engelmann et al. 1990).

For the first time since HEAO-3, several satellite or balloon-borne experiments (see ICRC 2007 reporter's talk Blasi 2008) have acquired higher quality data in the same energy range or covered a scarcely explored range (in terms of energy, 1 TeV/n–PeV/n, or in terms of nucleus): from the balloon-borne side, the ATIC collaboration has pre-

sented the B/C ratio at 0.5–50 GeV/n (Panov et al. 2007), and for H to Fe fluxes at 100 GeV–100 TeV (Panov et al. 2006). At higher energy, two long-duration balloon flights will soon provide spectra for $Z=1-30$ nuclei. The TRACER collaboration has published spectra for oxygen up to iron in the GeV/n–TeV/n range (Boyle et al. 2007; Ave et al. 2008). A second long-duration flight took place in summer 2006, during which the instrument was designed to have a wider dynamic-range capability and to measure lighter B, C, and N elements. The CREAM experiment (Seo et al. 2004) flew a cumulative duration of 70 days in December 2004 and December 2005 (Seo et al. 2006, and preliminary results in Marrocchesi et al. 2006 and Wakely et al. 2006), and again in December 2007. A fourth flight was scheduled for December 2008¹. Exciting data will arrive from the PAMELA satellite (Picozza et al. 2007), which was successfully launched in June 2006 (Casolino et al. 2008).

With this wealth of new data, it is relevant to question the method used to extract the propagation parame-

ters. The value of these parameters is important to many theoretical and astrophysical questions, because they are linked, amongst others, to the transport in turbulent magnetic fields, sources of CRs, and γ -ray diffuse emission (see Strong et al. 2007 for a recent review and references). It also proves to be crucial for indirect dark-matter detection studies (e.g., Donato et al. 2004, and Delahaye et al. 2008). The usage in the past has been based mostly on a manual or semi-automated—hence partial—coverage of the parameter space (e.g., Webber et al. 1992, Strong & Moskalenko 1998, and Jones et al. 2001). More complete scans were performed in Maurin et al. (2001, 2002), and Lionetto et al. (2005), although in an inefficient manner: the addition of a single new free parameter (as completed for example in Maurin et al. 2002 compared to Maurin et al. 2001) remains prohibitive in terms of computing time. To remedy these shortcomings, we propose to use the Markov Chain Monte Carlo (MCMC) algorithm, which is widely used in cosmological parameter estimates (e.g., Christensen et al. 2001, Lewis & Bridle 2002, and Dunkley et al. 2005). One goal of the paper is to confirm whether the MCMC algorithm can provide similar benefits in CR physics.

The analysis is performed in the framework of the Leaky-Box Model (LBM), a simple and widely used propagation model. This model contains most of the CR phenomenology and is well adapted to a first implementation of the MCMC tool. In Sect. 2, we highlight the appropriateness of the MCMC compared to other algorithms used in the field. In Sect. 3, the MCMC algorithm is presented. In Sect. 4, this algorithm is implemented in the LBM. In Sect. 5, we discuss the MCMC advantages and effectiveness in the field of CR physics, and present results for the LBM. We present our conclusions in Sect. 6. Application of the MCMC technique to a more up-to-date modelling, such as diffusion models, is left to a forthcoming paper.

2. Link between the MCMC, the CR data, and the model parameters

Various models describe the propagation of CRs in the interstellar medium (Webber et al. 1992; Bloemen et al. 1993; Strong & Moskalenko 1998; Maurin et al. 2001; Berezhko et al. 2003; Shibata et al. 2006; Evoli et al. 2008). Each model is based on his own specific geometry and has its own set of parameters, characterising the Galaxy properties. The MCMC approach aims to study quantitatively how the existing (or future) CR measurements can constrain these models or, equivalently, how in such models the set of parameters (and their uncertainties) can be inferred from the data.

In practice, a given set of parameters in a propagation model implies, e.g., a given B/C ratio. The model parameters are constrained such as to reproduce the measured ratio. The standard practice used to be an eye inspection of the goodness of fit to the data. This was replaced by the χ^2 analysis in recent papers: assuming the χ^2 statistics is applicable to the problem at stake, confidence intervals in these parameters can be extracted (see App. A).

The main drawback of this approach is the computing time required to extend the calculation of the χ^2 surface to a wider parameter space. This is known as the curse of dimensionality, due to the exponential increase in volume associated with adding extra dimensions to the parameter space, while the *good* regions of this space (for instance

where the model fits the data) only fill a tiny volume. This is where the MCMC approach, based on the Bayesian statistics, is superior to a grid approach. As in the grid approach, one end-product of the analysis is the χ^2 surface, but with a more efficient sampling of the region of interest. Moreover, as opposed to classical statistics, which is based on the construction of estimators of the parameters, Bayesian statistics assumes the unknown parameters to be random variables. As such, their full distribution—the so-called conditional probability-density function (PDF)—given some experimental data (and some prior density for these parameters, see below) can be generated.

To summarise, the MCMC algorithm provides the PDF of the model parameters, based on selected experimental data (e.g., B/C). The mean value and uncertainty in these parameters are by-products of the PDF. The MCMC enables the enlargement of the parameter space at a minimal computing time cost (although the MCMC and Metropolis-Hastings algorithms used here are not the most efficient one). The technicalities of the MCMC are briefly described below. The reader is referred to Neal (1993) and MacKay (2003) for a more substantial coverage of the subject.

3. Markov Chain Monte Carlo (MCMC)

Considering a model depending on m parameters

$$\boldsymbol{\theta} \equiv \{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(m)}\}, \quad (1)$$

we aim to determine the conditional PDF of the parameters given the data, $P(\boldsymbol{\theta}|\text{data})$. This so-called *posterior* probability quantifies the change in the degree of belief one can have in the m parameters of the model in the light of the data. Applied to the parameter inference, Bayes theorem is

$$P(\boldsymbol{\theta}|\text{data}) = \frac{P(\text{data}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})}{P(\text{data})}, \quad (2)$$

where $P(\text{data})$ is the data probability (the latter does not depend on the parameters and hence, can be considered to be a normalisation factor). This theorem links the posterior probability to the likelihood of the data $\mathcal{L}(\boldsymbol{\theta}) \equiv P(\text{data}|\boldsymbol{\theta})$ and the so-called *prior* probability, $P(\boldsymbol{\theta})$, indicating the degree of belief one has *before* observing the data. To extract information about a single parameter, $\theta^{(\alpha)}$, the posterior density is integrated over all other parameters $\theta^{(k \neq \alpha)}$ in a procedure called marginalisation. Finally, by integrating the individual posterior PDF further, we are able to determine the expectation value, confidence level, or higher order mode of the parameter $\theta^{(\alpha)}$. This illustrates the technical difficulty of Bayesian parameter estimates: determining the individual posterior PDF requires a high-dimensional integration of the overall posterior density. Thus, an efficient sampling method for the posterior PDF is mandatory. For models of more than a few parameters, regular grid-sampling approaches are not applicable and statistical techniques are required (Cowan 1997).

Among these techniques, MCMC algorithms have been fully tried and tested for Bayesian parameter inference (MacKay 2003; Neal 1993). MCMC methods explore any *target* distribution given by a vector of parameters $p(\boldsymbol{\theta})$, by generating a sequence of n points (hereafter a chain)

$$\{\boldsymbol{\theta}_i\}_{i=1,\dots,n} \equiv \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_n\}. \quad (3)$$

Each θ_i is a vector of m components [as defined in Eq. (1)]. In addition, the chain is Markovian in the sense that the distribution of θ_{n+1} is influenced entirely by the value of θ_n . MCMC algorithms are developed so that the time spent by the Markov chain in a region of the parameter space is proportional to the target PDF value in this region. Hence, from such a chain, one can obtain an independent sampling of the PDF. The target PDF as well as all marginalised PDF are estimated by counting the number of samples within the related region of parameter space.

Below, we provide a brief introduction to an MCMC using the Metropolis-Hastings algorithm (see Neal 1993 and MacKay 2003, chapter 29 for further details and references).

3.1. The algorithm

The prescription that we use to generate the Markov chains from the unknown target distribution is the so-called Metropolis-Hastings algorithm. The Markov chain increases by jumping from the current point in the parameter space θ_i to the following θ_{i+1} . As said before, the PDF of the new point only depends on the current point, i.e. $\mathcal{T}(\theta_{i+1}|\theta_1, \dots, \theta_i) = \mathcal{T}(\theta_{i+1}|\theta_i)$. This quantity defines the transition probability for state θ_{i+1} from the state θ_i . The Metropolis-Hastings algorithm specifies \mathcal{T} to ensure that the stationary distribution of the chain asymptotically tends to the target PDF one wishes to sample from.

At each step i (corresponding to a state θ_i), a trial state θ_{trial} is generated from a *proposal* density $q(\theta_{\text{trial}}|\theta_i)$. This proposal density is chosen so that samples can be easily generated (e.g., a Gaussian distribution centred on the current state). The state θ_{trial} is accepted or rejected depending on the following criterion. By forming the quantity

$$a(\theta_{\text{trial}}|\theta_i) = \min \left(1, \frac{p(\theta_{\text{trial}}) q(\theta_i|\theta_{\text{trial}})}{p(\theta_i) q(\theta_{\text{trial}}|\theta_i)} \right), \quad (4)$$

the trial state is accepted as a new state with a probability a (rejected with probability $1 - a$). The transition probability is then

$$\mathcal{T}(\theta_{i+1}|\theta_i) = a(\theta_{\text{trial}}|\theta_i) q(\theta_{\text{trial}}|\theta_i). \quad (5)$$

If accepted, $\theta_{i+1} = \theta_{\text{trial}}$, whereas if rejected, the new state is equivalent to the current state, $\theta_{i+1} = \theta_i$. This criterion ensures that once at its equilibrium, the chain samples the target distribution $p(\theta)$. If the proposal density $q(\theta_{\text{trial}}|\theta_i)$ is chosen to be symmetric, it cancels out in the expression of the acceptance probability, which becomes:

$$a = \min \left(1, \frac{p(\theta_{\text{trial}})}{p(\theta_i)} \right). \quad (6)$$

We note that the process requires only evaluations of ratios of the target PDF. This is a major virtue of this algorithm, in particular for Bayesian applications, in which the normalisation factor in Eq. (2), $P(\text{data}) = \int P(\text{data}|\theta) \cdot P(\theta) d\theta$ is often extremely difficult to compute. Hence, the ratio of the target PDF, i.e. the posterior of the parameter for our problem, can be calculated directly from the likelihood of the data and the priors.

3.2. Chain analysis

The chain analysis refers to the study of several properties of the chains. The following quantities are inspected in order to convert the chains in PDFs.

Burn-in length The burn-in describes the practice of removing some iterations at the beginning of the chain to eliminate the *transient* time needed to reach the equilibrium or stationary distribution, i.e., to *forget* the starting point. The burn-in length b is defined to be the number of first samples $\{\theta_i\}_{i=1, \dots, b}$ of the chain that must be discarded. The stationary distribution is reached when the chain enters the most probable parameter region corresponding to the region where the target function is close to its maximal value. To estimate b , the following criterion is used: we define $p_{1/2}$ to be the median of the target function distribution obtained from the entire chain of N samples. The burn-in length b corresponds to the first sample θ_b , for which $p(\theta_b) > p_{1/2}$ (see App. C for an illustration).

Correlation length By construction [see Eq. (5)], each step of the chain depends on the previous one, which ensures that the steps of the chain are correlated. We can obtain independent samples by thinning the chain, i.e. by selecting only a fraction of the steps with a periodicity chosen to derive uncorrelated samples. This period is estimated by computing the autocorrelation functions for each parameter. For a parameter $\theta^{(\alpha)}$ ($\alpha = 1, \dots, m$), the autocorrelation function is given by

$$c_j^{(\alpha)} = \frac{E \left[\theta_i^{(\alpha)} \theta_{j+i}^{(\alpha)} \right] - \left(E \left[\theta_i^{(\alpha)} \right] \right)^2}{E \left[\left(\theta_i^{(\alpha)} \right)^2 \right]}, \quad (7)$$

which we calculate with the Fast Fourier Transformation (FFT). The correlation length $l^{(\alpha)}$ for the α -th parameter is defined as the smallest j for which $c_j^{(\alpha)} < 1/2$, i.e. the values $\theta_i^{(\alpha)}$ and $\theta_{i+j}^{(\alpha)}$ of the chain that are considered to be uncorrelated. The correlation length l for the chain, for all parameters, is defined to be

$$l \equiv \max_{\alpha=1, \dots, m} l^{(\alpha)}, \quad (8)$$

which is used as the period of the thinning (see App. C for an illustration).

Independent samples and acceptance The independent samples of the chain are chosen to be $\{\theta_i\}_{i=b+lk}$, where k is an integer. The number of independent samples N_{ind} is defined to be the fraction of steps remaining after discarding the burn-in steps and thinning the chain,

$$N_{\text{ind}} = \frac{N_{\text{tot}} - b}{l}. \quad (9)$$

The independent acceptance f_{ind} is the ratio of the number of independent samples N_{ind} to the total step number N_{tot} ,

$$f_{\text{ind}} = \frac{N_{\text{ind}}}{N_{\text{tot}}}. \quad (10)$$

3.3. Choice of the target and trial functions

3.3.1. Target function

As already said, we wish to sample the target function $p(\theta) = P(\theta|\text{data})$. Using Eq. (2) and the fact that

the algorithm is insensitive to the normalisation factor, this amounts to sampling the product $P(\text{data}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})$. Assuming a flat prior $P(\boldsymbol{\theta}) = \text{cst}$, the target distribution reduces to

$$p(\boldsymbol{\theta}) = P(\text{data}|\boldsymbol{\theta}) \equiv \mathcal{L}(\boldsymbol{\theta}), \quad (11)$$

and here, the likelihood function is taken to be

$$\mathcal{L}(\boldsymbol{\theta}) = \exp\left(-\frac{\chi^2(\boldsymbol{\theta})}{2}\right). \quad (12)$$

The $\chi^2(\boldsymbol{\theta})$ function for n_{data} data is

$$\chi^2(\boldsymbol{\theta}) = \sum_{k=1}^{n_{\text{data}}} \frac{(y_k^{\text{exp}} - y_k^{\text{theo}}(\boldsymbol{\theta}))^2}{\sigma_k^2}, \quad (13)$$

where y_k^{exp} is the measured value, y_k^{theo} is the hypothesised value for both a certain model and the parameters $\boldsymbol{\theta}$, and σ_k is the known variance of the measurement. For example, y_k^{exp} and y_k^{theo} represent the measured and calculated B/C ratios.

The link between the target function, i.e., the posterior PDF of the parameters, and the experimental data is established with the help of Eqs (11) to (13). This link guarantees the proper sampling of the parameter space using Markov chains, which spend more time in more relevant regions of parameter space, as described above.

3.3.2. Trial function

Despite the effectiveness of the Metropolis-Hastings algorithm, to optimise the efficiency of the MCMC and minimise the number of chains to be processed, trial functions should be as close as possible to the true distributions. We use a sequence of three trial functions to explore the parameter space. The first step is a coarse determination of the parameter PDF. This allows us to calculate the covariance matrix leading to a better coverage of parameter space, provided that the target PDF is sufficiently close to being an N-dimensional Gaussian. The last step takes advantage of a binary-space partitioning (BSP) algorithm.

Gaussian step For the first iteration, the proposal density $q(\boldsymbol{\theta}_{\text{trial}}, \boldsymbol{\theta}_i)$, required to obtain the trial value $\boldsymbol{\theta}_{\text{trial}}$ from $\boldsymbol{\theta}_i$ is written as

$$q(\boldsymbol{\theta}_{\text{trial}}, \boldsymbol{\theta}_i) \propto \prod_{\alpha=1, \dots, m} \exp\left(-\frac{1}{2} \frac{(\theta_{\text{trial}}^{(\alpha)} - \theta_i^{(\alpha)})^2}{\sigma_\alpha^2}\right). \quad (14)$$

These represent m independent Gaussian distributions centred on $\boldsymbol{\theta}_i$. The distribution is symmetric, so that the acceptance probability a follows Eq. (6). The variance σ_α^2 for each parameter α is to be specified. Each parameter $\theta_{\text{trial}}^{(\alpha)}$ is hence calculated to be

$$\theta_{\text{trial}}^{(\alpha)} = \theta_i^{(\alpha)} + \sigma_\alpha \cdot x,$$

where x is a random number obeying a Gaussian distribution centred on zero with unit variance.

It is important to choose an optimal width σ_α to sample properly the posterior (target) distribution. If the width is too large, as soon as the chain reaches a region of high probability, most of the trial parameters fall into a region of low

probability and are rejected, leading to a low acceptance and a long correlation length. Conversely, for too small a width, the chain will take a longer time to reach the interesting regions. Eventually, even if the chain reaches these regions of high acceptance, only a partial coverage of the PDF support will be sampled (also leading to a long correlation length).

In practice, we first define σ_α ($\alpha = 1, \dots, m$) equal to the expected range of the parameter. In a subsequent iteration, σ_α is set to be $2\sqrt{2 \ln 2} \approx 2.3$ times $\sigma_\alpha^{\text{calc}}$, i.e. the FWHM of the PDF obtained with the first iteration. The result is actually insensitive to the numerical factor used.

Covariance matrix The proposal density is taken to be an N-dimensional Gaussian of covariance matrix V

$$q(\boldsymbol{\theta}_{\text{trial}}, \boldsymbol{\theta}_i) \propto \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_{\text{trial}} - \boldsymbol{\theta}_i)^T V^{-1}(\boldsymbol{\theta}_{\text{trial}} - \boldsymbol{\theta}_i)\right). \quad (15)$$

The covariance matrix V is symmetric and diagonalisable (D is a diagonal matrix of eigenvalues and P represents the change in the coordinate matrix),

$$V = P^T D P,$$

and where again Eq. (6) holds. The parameters $\boldsymbol{\theta}_{\text{trial}}$ are hence found to be

$$\boldsymbol{\theta}_{\text{trial}} = \boldsymbol{\theta}_i + P^T D \boldsymbol{x},$$

where \boldsymbol{x} is a vector of m random numbers following a Gaussian distribution centred on zero and with unit variance.

The covariance matrix V is estimated, e.g., from a previous iteration using the Gaussian step. The advantage of this trial function with respect to the previous one is that it takes account of the possible correlations between the m parameters of the model.

Binary Space Partitioning (BSP) A third method was developed to define a proposal density for which the results of the Gaussian step or the covariance matrix iterations are used to subdivide the parameter space into boxes, in each of which a given probability is affected.

The partitioning of the parameter space can be organised using a binary-tree data structure known as a binary-space partitioning tree (de Berg et al. 2000). The root node of the tree is the m -dimensional box corresponding to the entire parameter space. The binary-space partitioning is then performed by dividing each box recursively into two child boxes if the partitioning satisfies the following requirement: a box is divided only if the number of independent samples contained in this box is higher than a certain number (here we used a maximum of between 3% and 0.1% of the total number of independent samples). When a box has to be divided, the division is made along the longer side of the box (the box-side lengths are defined relative to the root-box sides). For each end node (i.e. node without any children), a probability, defined as the fraction of the number of independent samples in the box to their total number, is assigned. For empty boxes, a minimum probability is assigned and all the probabilities are renormalised so that the sum of all end-node probabilities equals 1.

The proposal density $q(\boldsymbol{\theta}_{\text{trial}})$ is then defined, in each end-node box, as a uniform function equal to the assigned

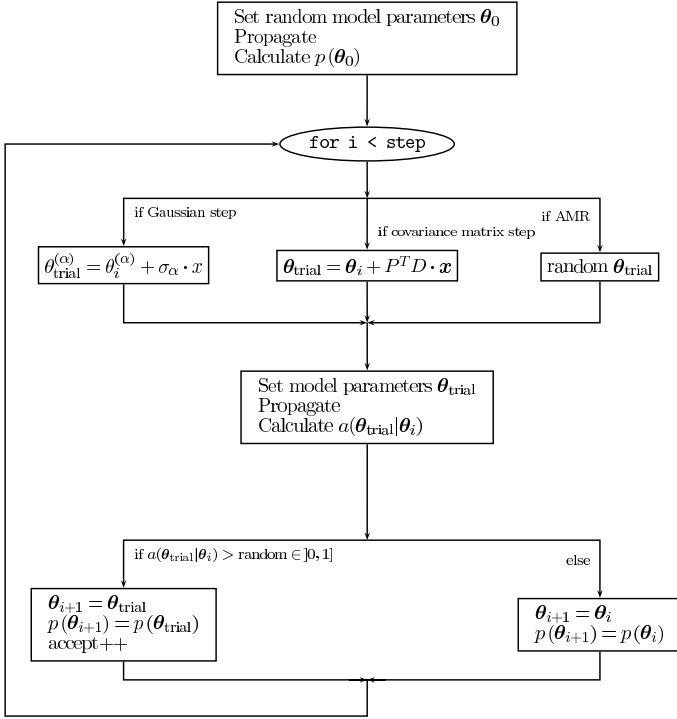


Fig. 1. Flow chart of the implemented MCMC algorithm: θ_i is a vector [Eq. (1)] of the $\alpha = 1, \dots, m$ free parameters of the model, evaluated at each step i , and $p(\theta_i)$ is the target function given by Eq. (11). See text for details.

probability. The sampling of this proposal density is simple and efficient: an end node is chosen with the assigned probability and the trial parameters are chosen uniformly in the corresponding box. In comparison to the other two proposal densities, this proposal density based on a BSP is asymmetric, because it is only dependent on the proposal state $q(\theta_{\text{trial}})$. Hence, Eq. (4) must be used.

4. Implementation in the propagation model

The MCMC with the three above methods are implemented in the USINE package², which computes the propagation of Galactic CR nuclei and anti-nuclei for several propagation models (LBM, 1D and 2D diffusion models). The reader is referred to Maurin et al. (2001) for a detailed description for the nuclear parameters (fragmentation and absorption cross-sections), energy losses (ionisation and Coulomb), and solar modulation (force-field) used.

We briefly describe how the MCMC algorithm is implemented in the propagation part (Sect. 4.1), using a LBM—the procedure would be similar for any other model. The LBM and its parameters are briefly discussed (Sect. 4.2) as well as the input spectrum parameters (Sect. 4.3). Additional information about the data are gathered in App. B.

4.1. Flow chart

A flow chart of the Metropolis-Hastings MCMC algorithm used in the context of GCRs is given in Fig. 1. To sum-

marise, the initial values of the propagation parameters θ_0 are chosen randomly in their expected range to crank up each Markov chain. The interstellar (IS) CR fluxes are then calculated for this set of parameters (see, e.g., Fig. 1 in Maurin et al. 2002 for further details of the propagation steps). The IS flux is modulated with the force-field approximation and the resulting top-of-atmosphere (TOA) spectrum is compared with the data, which allows us to calculate the χ^2 value [Eq. (13)], hence the likelihood [Eq. (12)]. This likelihood (in practice the log-likelihood) is used to compute the acceptance probability [Eq. (4)] of the trial vector of parameters θ_{trial} (as generated by one of the three trial functions described in Sect. 3.3.2). Whether the trial vector is accepted or rejected implies whether $\theta_1 = \theta_{\text{trial}}$ or $\theta_1 = \theta_0$. This procedure is repeated for the N steps of the chain. Obviously, when $\theta_{i+1} = \theta_i$, the propagation step does not need to be repeated. Because of the nature of the MCMC algorithm, several chains can be executed in parallel. Once completed, these chains are analysed (see Sect. 3.2)—discarding the first step belonging to the burning length and thinned according to the correlation length l [Eq. (8)]—and combined to recover the desired posterior PDF $P(\theta|\text{data})$.

In this procedure, the user must decide i) the data to be used, ii) the observable to retain in calculating the likelihood, and iii) the number of free parameters m (of the vector θ) for which we seek the posterior PDF.

4.2. Leaky-Box Model (LBM)

The LBM assumes that all CR species are confined within the Galaxy with an escape rate that equals N/τ_{esc} , where the escape time τ_{esc} is rigidity-dependent, and is written as $\tau_{\text{esc}}(R)$. This escape time has two origins. First, CRs can leak out the confinement volume and leave the Galaxy. Second, they can be destructed by spallation on interstellar matter nuclei. This latter effect is parameterised by the grammage x (usually expressed in g cm^{-2}), defined as the column density of interstellar matter encountered by a path followed by a CR. The CRs that reach Earth have followed different paths, and can therefore be described by a grammage distribution $N(x) \equiv dN/dx$. The LBM assumes that

$$N(x) \propto \exp^{-\lambda_{\text{esc}}(R)x}, \quad (16)$$

where the mean grammage $\lambda_{\text{esc}}(R) = \langle x \rangle$ is related to the mass m , velocity v and escape time $\tau_{\text{esc}}(R)$ by means of $\lambda_{\text{esc}}(R) = \bar{m}nv\tau_{\text{esc}}(R)$.

The function $\lambda_{\text{esc}}(R)$ determines the amount of spallations experienced by a primary species, and thus determines the secondary-to-primary ratios, for instance B/C. From an experimentalist point of view, $\lambda_{\text{esc}}(R)$ is a quantity that can be inferred from measurements of nuclei abundance ratios. The grammage $\lambda_{\text{esc}}(R)$ is known to provide an effective description of diffusion models (Berezinskii et al. 1990): it can be related to the efficiency of confinement (which is determined by the diffusion coefficient and to both the size and geometry of the diffusion volume), spallative destruction (which tends to shorten the average lifetime of a CR and thus lower λ_{esc}), and a mixture of other processes (such as convection, energy gain, and losses).

In this paper, we compute the fluxes in the framework of the LBM with minimal reacceleration by the interstellar turbulence, as described in Osborne & Ptuskin (1988) and

² A public version will be released soon (Maurin, in preparation).

Seo & Ptuskin (1994). The grammage $\lambda_{\text{esc}}(R)$ is parameterised as

$$\lambda_{\text{esc}}(R) = \begin{cases} \lambda_0 \beta R_0^{-(\delta-\delta_0)} R^{-\delta_0} & \text{when } R < R_0, \\ \lambda_0 \beta R^{-\delta} & \text{otherwise;} \end{cases} \quad (17)$$

where we allow for a break, i.e. a different slope below and above a critical rigidity R_0 . The standard form used in the literature is recovered by setting $\delta_0 = 0$. For the entire set of n nuclei, a series of n equations (see Maurin et al. 2001 for more details) for the differential densities $N^{j=1,\dots,n}$ are solved at a given kinetic energy per nucleon $E_{k/n}$ (E is the total energy), i.e.

$$A^j N^j(E_{k/n}) + \frac{d}{dE} \left(B^j N^j - C^j \frac{dN^j}{dE} \right) = S^j(E_{k/n}). \quad (18)$$

In this equation, the r.h.s. term is the source term that takes into account the primary contribution (see Sect. 4.3), the spallative secondary contribution from all nuclei k heavier than j , and the β -decay of radioactive nuclei into j . The first energy-dependent factor A^j is given by

$$A^j = \frac{1}{\tau_{\text{esc}}} + \sum_{ISM=H,He} n_{\text{ISM}} v^j \sigma_{\text{inel}}^{j+\text{ISM}} + \frac{1}{\tau_{\beta}^j}.$$

The two other terms correspond to energy losses and first-order reacceleration for B^j and to second-order reacceleration for C^j . Following Osborne & Ptuskin (1988) and Seo & Ptuskin (1994),

$$B = \left\langle \frac{dE}{dt} \right\rangle_{\text{ion, coul.}} + (1+\beta^2) \beta^2 E K_{pp} \quad \text{and} \quad C = \beta^4 E^2 K_{pp},$$

where

$$K_{pp} = \frac{4}{3} \mathcal{V}_a^2 \frac{\tau_{\text{esc}}}{\delta(4-\delta^2)(4-\delta)}. \quad (19)$$

The strength of the reacceleration is mediated by the pseudo Alfvénic speed \mathcal{V}_a of the scatterers in units of $\text{km s}^{-1} \text{ kpc}^{-1}$. This is related to a true speed given in a diffusion model with a thin disk h and a diffusive halo L by means of $\mathcal{V}_a = V_a \times (hL)^{-1/2}$ (Seo & Ptuskin 1994). Assuming typical values of $h = 0.1 \text{ kpc}$ and $L = 10 \text{ kpc}$, the value of \mathcal{V}_a can be directly transposed and compared to a true speed V_a , as obtained in diffusion models.

To summarise, our LBM with reacceleration may involve up to five free parameters, i.e. the normalisation λ_0 , the slopes δ_0 and δ below or above the cut-off rigidity R_0 , and a pseudo-Alfvén velocity \mathcal{V}_a related to the reacceleration strength.

4.3. Source spectra

We assume that the primary source spectrum $Q_j(E)$ for each nuclear species j is given by ($\beta = v/c$)

$$Q_j(E) \equiv dQ_j/dE = q_j \beta^{\eta_j} R^{-\alpha_j}, \quad (20)$$

where q_j is the source abundance, α_j is the slope of the species j , and the term β^{η_j} manifests our ignorance about the low-energy spectral shape. We further assume that $\alpha_j \equiv \alpha$ for all j , and unless stated otherwise, $\eta_j \equiv \eta = -1$ in order to recover $dQ/dp \propto p^{-\alpha}$, as obtained from acceleration models (e.g., Jones 1994). The constraints existing on η are explored in Sect. 5.3.

The pattern of the source abundances observed in the cosmic radiation differs from that of the solar system. This is due to a segregation mechanism during the acceleration stage. Two hypotheses are disputed in the literature: one is based on the CR composition controlled by volatility and mass-to-charge ratio (Meyer et al. 1997; Ellison et al. 1997), and the other one is based on the first ionisation potential (FIP) of nuclei (e.g., Cassé & Goret 1973). In this work, for each configuration, the source abundances are initialised to the product of the solar system abundances (Lodders 2003), and the value of the FIP taken from Binns et al. (1989). The final fluxes are obtained by an iterative calculation of the propagated fluxes, rescaling the element abundances—keeping fixed the relative isotopic abundances—to match experimental data at each step until convergence is reached (see Fig. 1 in Maurin et al. 2002 for further details). The result is thus insensitive to the input values (more details about the procedure are given in App. B.1).

The measurement of all propagated isotopic fluxes should characterise all source spectra parameters completely, i.e. the q_j and α_j parameters should be free. However, only element fluxes are available, which motivates the above *rescaling* approach. In Sect. 5.3, a few calculations are undertaken to determine self-consistently, along with the propagation parameters, i) α and η , and ii) the source abundances for the primary species C, O, and the mixed N elements (the main contributors to the boron flux).

5. Results

We first examine the relative merits of four different parameterisations of the LBM, and determine the statistical significance of adding more parameters. These models correspond to $\{\theta^\alpha\}_{\alpha=1,\dots,m \leq 5}$ with

- Model I = $\{\lambda_0, R_0, \delta\}$, i.e. no reacceleration ($\mathcal{V}_a = 0$) and no break in the spectral index ($\delta_0 = 0$).
- Model II = $\{\lambda_0, \delta, \mathcal{V}_a\}$, i.e. no critical rigidity ($R_0 = 0$) and no break in the spectral index ($\delta_0 = 0$).
- Model III = $\{\lambda_0, R_0, \delta, \mathcal{V}_a\}$, i.e. no break in the spectral index ($\delta_0 = 0$).
- Model IV = $\{\lambda_0, R_0, \delta_0, \delta, \mathcal{V}_a\}$.

Various subsets of B/C data are used to investigate whether old data are useful or just add confusion to the PDF determination. We note in Sect. 5.2 that no useful constraint can be drawn from \bar{p} data alone.

We also consider additional free parameters (Sect. 5.3) related to the source spectra, for a self-consistent determination of the propagation and source properties. Since we show that a break in the slope (Model IV) is not required by current data, we focus on Model III (for the description of the propagation parameters), defining:

- Model III+1 = $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha\}$, where the source slope α is a free parameter.
- Model III+2 = $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, \eta\}$, where both the source slope α and the exponent η [of β , see Eq. (20)] are free parameters.
- Model III+4 = $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, q_C, q_N, q_O\}$, where the abundances q_i of the most significantly contributing elements are also free parameters.
- Model III+5 = $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, \eta, q_C, q_N, q_O\}$.

This allows us to investigate the correlations between parameters further and into potential biases in the propagation parameter determination.

More details about the practical use of the trial functions can be found in App. C. In particular, the sequential use of the three sampling methods (Gaussian step, covariance matrix step, and then binary-space partitioning) is found to be the most efficient: all results presented hereafter are based on this sequence.

5.1. Fitting the B/C ratio

5.1.1. HEAO-3 data alone

We first constrain the model parameters with HEAO-3 data only (Engelmann et al. 1990). These data are the most precise data available at the present day for the stable nuclei ratio B/C of energy between 0.62 to 35 GeV/n.

The results for the models I, II, and III are presented in Figs. C.2, 2 top, and 2 bottom. The inner and outer contours are taken to be regions containing 68% and 95% of the PDF respectively (see App. A.1). The first observation one can make for the LBM without reacceleration (Model I, Fig. C.2), is that the marginal distributions of the three LBM parameters are mostly Gaussian. The tail for small values of R_0 is due to this parameter being constrained by low-energy data (< 1 GeV/n): there are no HEAO-3 data at low energy, so all R_0 values below 3 GV are equiprobable (this remains true for Model III).

As seen in Fig. 2, a more complicated shape for the different parameters is found for Model II (top panel), and even more so for Model III (bottom panel). This induces a longer correlation length (1.5 and 6.9 steps instead of 1 step) and hence reduces the efficiency of the MCMC (75% for model II and 17% for model III). Physically, the correlation between the parameters, as seen most clearly in Fig. 2 (bottom), is understood as follows. First, λ_0 , R_0 , and δ are positively correlated. This originates in the low-energy relation $\lambda_{\text{esc}} \propto \lambda_0 R_0^{-\delta}$, which should remain approximately constant to reproduce the bulk of the data at GeV/n energy. Hence, if R_0 or δ is increased, λ_0 also increases to balance the product. On the other hand, \mathcal{V}_a is negatively correlated with δ (and hence with all the parameters): this is the standard result that to reach smaller δ (for instance to reach a Kolmogorov spectrum), more reacceleration is required. This can also be seen from Eq. (19), where at constant τ_{esc} , $K_{pp} \propto \mathcal{V}_a^2 / f(\delta)$, where f is a decreasing function of δ : hence, if δ decreases, $f(\delta)$ increases, and \mathcal{V}_a then has to increase to retain the balance.

The values for the maximum of the PDF for the propagation parameters along with their 68% confidence intervals (see App. A) are listed in Table 1. The values obtained for our Model I are in fair agreement with those derived by Webber et al. (1998), who found $\{\lambda_0, R_0, \delta\} = \{38.27, 3.6, 0.7\}$. The difference for λ_0 could be related to the fact that Webber et al. (1998) rely on a mere eye inspection to extract the best-fit solution or/and use a different set of data. For example, comparing Model I with a combination of HEAO-3 and low-energy data (ACE+Voyager 1 & 2+IMP7-8, see Sect. 5.1.2) leads to $\{\lambda_0, R_0, \delta\} = \{52, 5.3, 0.69\}$, slightly changing the values of the Model I preferred parameters (compared to the first line of Table 1).

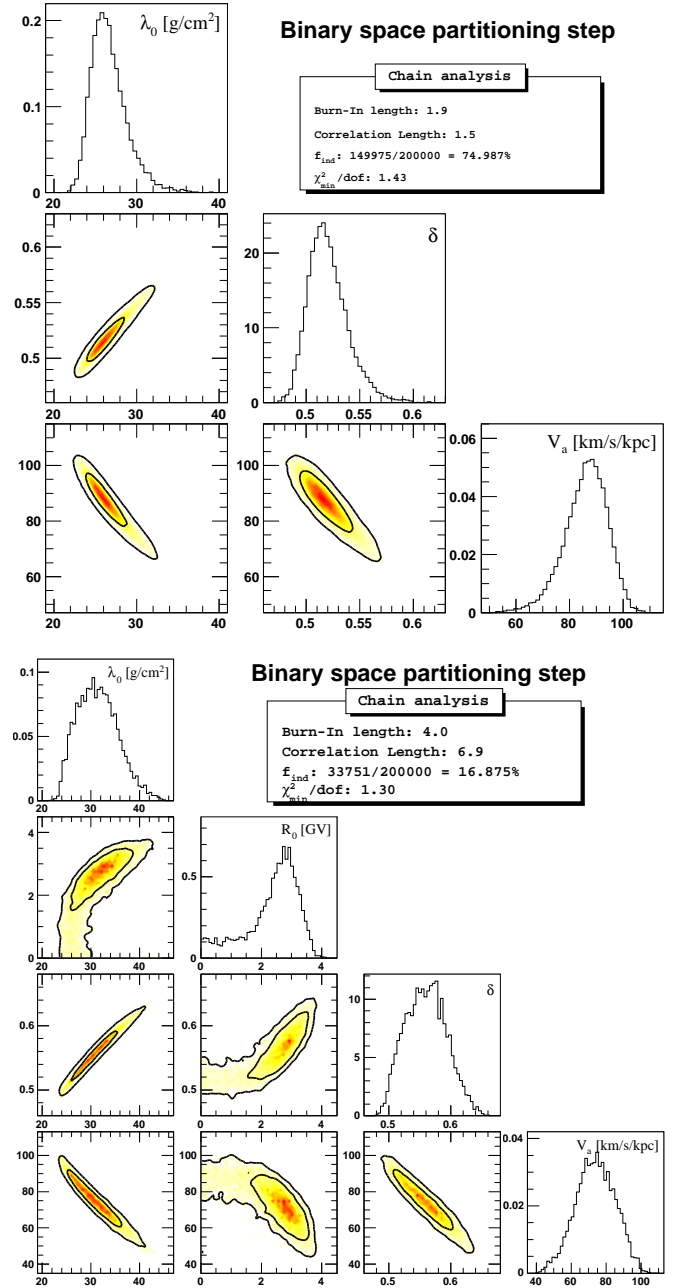


Fig. 2. Posterior distributions for Model II (top) and Model III (bottom) using HEAO-3 data only. For more details, refer to caption of Fig. C.2.

The reacceleration mechanism was invoked in the literature to decrease the spectral index δ toward its preferred value of 1/3 given by a Kolmogorov spectrum of turbulence. In Table 1, the estimated propagation parameter values for the models II and III are indeed slightly smaller than for Model I, but the Kolmogorov spectral index is excluded for all of these three cases (using HEAO-3 data only). This result agrees with the findings of Maurin et al. (2001), in which a more realistic two-dimensional diffusion model with reacceleration and convection was used. We note that the values for $\mathcal{V}_a \sim 80 \text{ km s}^{-1} \text{ kpc}^{-1}$, should lead to a true speed $V_a = \mathcal{V}_a \times \sqrt{hL} \sim 80 \text{ km s}^{-1}$ in a diffusion model for which the thin disk half-height is $h = 0.1 \text{ kpc}$ and the halo size is $L = 10 \text{ kpc}$: this is consistent with values found in Maurin et al. (2002).

Model	λ_0 g cm ⁻²	R_0 GV	δ	\mathcal{V}_a km s ⁻¹ kpc ⁻¹	χ^2_{\min}/dof
I	54 ⁺² ₋₂	4.2 ^{+0.3} _{-0.9}	0.70 ^{+0.01} _{-0.01}	-	3.35
II	26 ⁺² ₋₂	-	0.52 ^{+0.02} _{-0.02}	88 ⁺⁶ ₋₁₁	1.43
III	30 ⁺⁵ ₋₄	2.8 ^{+0.6} _{-0.8}	0.58 ^{+0.01} _{-0.06}	75 ⁺¹⁰ ₋₁₃	1.30

Table 1. Most probable values of the propagation parameters (after marginalising over the other parameters) for models I, II, and III using HEAO-3 alone (14 data points) and the B/C constraint. The uncertainty in the parameters correspond to 68% CL of the marginalised PDF (see App. A). The last column shows the minimum χ^2/dof obtained for each model (the associated best-fit parameters are gathered in Table 3).

Model	λ_0 g cm ⁻²	R_0 GV	δ	\mathcal{V}_a km s ⁻¹ kpc ⁻¹	χ^2_{\min}/dof
Dataset					
III-A	30 ⁺⁵ ₋₄	2.8 ^{+0.6} _{-0.8}	0.58 ^{+0.01} _{-0.06}	75 ⁺¹⁰ ₋₁₃	1.30
III-B	26 ⁺² ₋₃	2.6 ^{+0.4} _{-0.7}	0.53 ^{+0.02} _{-0.03}	85 ⁺⁹ ₋₈	1.09
III-C	27 ⁺² ₋₂	2.6 ^{+0.4} _{-0.7}	0.53 ^{+0.02} _{-0.03}	86 ⁺⁹ ₋₅	1.06
III-D	26 ⁺² ₋₂	3.0 ^{+0.4} _{-0.5}	0.52 ^{+0.02} _{-0.02}	95 ⁺⁷ ₋₆	4.15
III-E	30 ⁺² ₋₂	3.7 ^{+0.2} _{-0.3}	0.57 ^{+0.01} _{-0.02}	88 ⁺³ ₋₆	6.08

Table 2. Same as in Table 1, but testing different data sets with Model III: A = HEAO-3 (14 data points), B = HEAO-3 + ACE (20 data points), C = HEAO-3 + ACE + Voyager 1 & 2 + IMP7-8 (22 data points), D = HEAO-3 + all low-energy data (30 data points), E = all B/C data (69 data points).

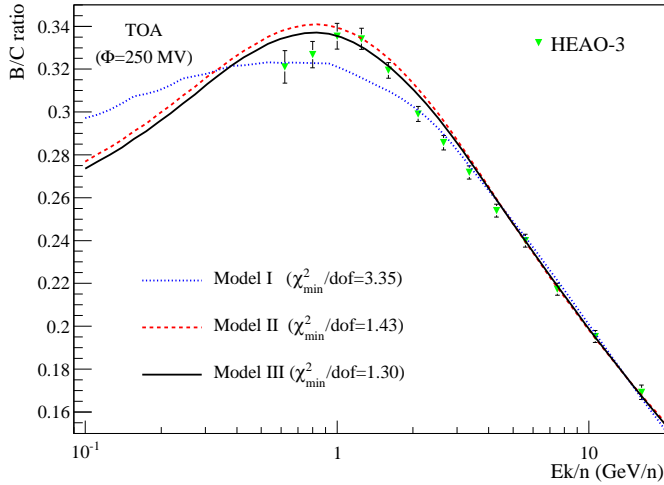


Fig. 3. Best-fit ratio for Model I (blue dotted), II (red dashed), and Model III (black solid) using the HEAO-3 data only (green symbols). The curves are modulated with $\Phi = 250$ GV. The corresponding best-fit parameters are gathered in Table 3.

The final column in Table 1 indicates, for each model, the best χ^2 value per degree of freedom, χ^2_{\min}/dof . This allows us to compare the relative merit of the models. LB models with reacceleration reproduce the HEAO-3 data more accurately (with χ^2/dof of 1.43 and 1.30 for the Models II and III respectively compared to $\chi^2/\text{dof} = 4.35$ for Model I). The best-fit model B/C fluxes are shown with the B/C HEAO-3 data modulated at $\Phi = 250$ MV in Fig. 3. Physically, the origin of a cutoff R_0 in λ_{esc} at low energy can be related to convection in diffusion models (Jones 1979). Hence, it is a distinct process as reacceleration. The fact that Model III performs more successfully than Model II implies that both processes are significant, as found in Maurin et al. (2001).

In the following, we no longer consider Model I and II, and inspect instead, the parameter dependence of Model III on the dataset selected.

5.1.2. Additional constraints from low-energy data

The actual data sets for the B/C ratio (see, e.g., Fig. 5) show a separation into two energy domains: the low-energy range extends from $\sim 10^{-2}$ GeV/n to ~ 1 GeV/n and the high-energy range goes from ~ 1 GeV/n to $\sim 10^2$ GeV/n. The spectral index δ is constrained by high-energy data,

e.g., the HEAO-3 data, and adding low-energy data allows us to more reliably constrain R_0 . We note that by fitting only the low-energy data, only the grammage crossed in very narrow energy domain would be constrained.

In a first step, we add only the ACE (CRIS) data (de Nolfo et al. 2006), which covers the energy range from $\sim 8 \cdot 10^{-2}$ GeV/n to $\sim 2 \cdot 10^{-1}$ GeV/n, and which is later referred to as dataset B (the dataset A being HEAO-3 data alone). The resulting posterior distributions are similar for the datasets B and A (B is not shown, but A is given in Fig. 2, bottom). Results for datasets A and B are completely consistent (first and second line of Table 2), but for the latter, propagation parameters are more tightly constrained and the fit is improved ($\chi^2_{\min}/\text{dof}=1.09$). The ACE (CRIS) data are compatible with $R_0 = 0$, but the preferred critical rigidity is 2.47 GV.

All other low-energy data (ISEE-3, Ulysses, IMP7-8, Voyager 1 & 2, ACE) are then included (dataset D). The resulting values of the propagation parameters are left unchanged. However, a major difference lies in the higher χ^2_{\min}/dof of 4.15, which reflects an inconsistency between the different low-energy data chosen for the MCMC. If the data point from the Ulysses experiment is excluded, χ^2_{\min}/dof decreases to a value of 2.26, and by excluding also the ISEE-3 data points (dataset C) it decreases further to 1.06 (see Table 2). Since the set of low-energy data have different modulation parameters, the difference in the results for the various data subsets becomes clearer after the data have been demodulated. The force-field approximation provides a simple analytical one-to-one correspondence between the modulated top-of-the-atmosphere (TOA) and the demodulated interstellar (IS) fluxes. For an isotope x , the IS and TOA energies per nucleon are related by $E_k^{\text{IS}} = E_k^{\text{TOA}} + \Phi$ ($\Phi = Z/A \times \phi$ is the modulation parameter), and the fluxes by (p_x is the momentum per nucleon of x)

$$\psi_x^{\text{IS}}(E_k^{\text{IS}}) = \left(\frac{p_x^{\text{IS}}}{p_x^{\text{TOA}}} \right)^2 \psi_x^{\text{TOA}}(E_k^{\text{TOA}} + Z/A \times \psi). \quad (21)$$

The B/C ratio results from a combination of various isotopes, and assuming the same Z/A for all isotopes, we find that

$$\left(\frac{B}{C} \right)^{\text{IS}}(E_k^{\text{IS}}) = \left(\frac{B}{C} \right)^{\text{TOA}}(E_k^{\text{TOA}} + Z/A \times \phi). \quad (22)$$

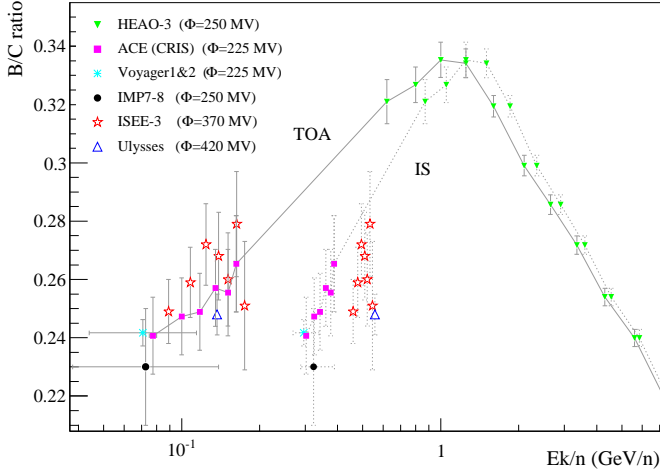


Fig. 4. HEAO-3 and ACE (CRIS) modulated (TOA, solid line) and demodulated (IS, dashed line) data points have been connected to guide the eye. Filled symbols (modulated and demodulated) correspond to HEAO-3, ACE (CRIS), IMP7-8 and Voyager 1 & 2. On the TOA curve, the empty red stars and the blue upper triangle correspond to ISEE-3 and Ulysses.

Model	λ_0^{best}	R_0^{best}	δ_0^{best}	δ^{best}	V_a^{best}	χ^2/dof
Data	g cm^{-2}	GV			$\text{km s}^{-1} \text{kpc}^{-1}$	
I-A	54.7	4.21	-	0.702	-	3.35
II-A	25.8	-	-	0.514	88.8	1.43
III-A	31.7	2.73	-	0.564	73.0	1.30
III-C	26.9	2.45	-	0.527	88.5	1.06
IV-C	32.7	2.38	-0.97	0.572	70.5	0.86

Table 3. Best-fit values (corresponding to χ_{min}^2) for B/C data (A=HEAO-3 data alone, C=HEAO-3+Voyager 1 & 2+ACE+IMP7-8).

The modulated and demodulated low-energy B/C data are shown in Fig. 4 (see caption for details). The ISEE-3 and Ulysses data points, as just underlined, are clearly inconsistent with other data. To be consistent, $\Phi = 200$ MV for ISEE-3 and $\Phi = 200$ MV for Ulysses would be required. Significant uncertainties $\Delta\Phi \sim 25 - 50$ GV are quoted in general, so that it is difficult to conclude whether there are systematics in the measurement or if the modulation quoted in the papers is inappropriate. Some experiments have also accumulated the signal for several years, periods during which the modulation changes. It is beyond the scope of this paper to discuss this issue further. Below, we discard both ISEE-3 and Ulysses data in selecting an homogeneous low-energy data set, which includes the most recent ACE (CRIS) data.

The resulting best-fit models, when taking low-energy data into account, are displayed in Fig. 5. The B/C best-fit ratio is displayed for Model III and the dataset A (red thin lines) and C (black thick lines). Model III-B (not shown) yields similar results to Model III-C. Solid and dashed lines correspond to the two modulations $\Phi = 250$ MV (HEAO-3 and IMP7-8) and $\Phi = 225$ MV respectively (ACE and Voyager 1 & 2). Although the fit from HEAO-3 alone provides a good match at low energy, adding ACE (CRIS) and

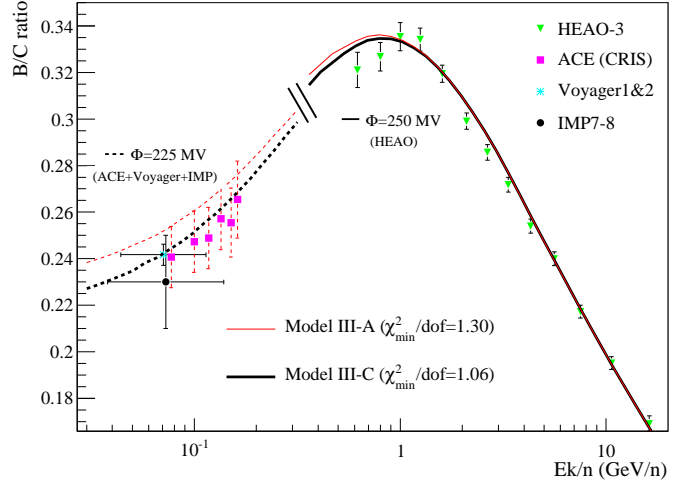


Fig. 5. Best-fit B/C flux (Model III) for datasets A (thin red curves) and C (thick black curves). Above 300 MeV/n, B/C is modulated to $\Phi = 250$ MV (solid lines) appropriate for HEAO-3 data, whereas below, it is modulated to $\Phi = 225$ MV (dashed lines). Model III-B, not shown, overlaps with III-C. The corresponding propagation values are gathered in Table 3.

Voyager 1 & 2 constraints slightly shifts all of the parameters to lower values.

In a final try, we take into account all available data (dataset E, final line of Table 2). Many data are clearly inconsistent with each other (see Fig. 8), but as for the low-energy case, although the $\chi_{\text{min}}^2/\text{dof}$ is worsened, the preferred values of the propagation parameters are not changed drastically (compare with datasets B, C, and D in Table 2). We await forthcoming data from CREAM, TRACER, and PAMELA to be able to confirm and refine the results for HEAO-3 data.

5.1.3. Model IV: break in the spectral index

We have already mentioned that the rigidity cut-off may be associated with the existence of a galactic wind in diffusion models. By allowing a break to occur in the spectral index of λ_{esc} [see Eq. (17)], we search for a deviations from a single power law ($\delta_0 = \delta$) or from the cut-off case ($\delta_0 = 0$).

Adding a new parameter δ_0 (Model IV) increases the correlation length of the MCMC, since R_0 and δ_0 are correlated [see Eq. (17)]. The acceptance f_{ind} [Eq. (9)] is hence extremely low. For Model IV-C (i.e. using dataset C, see Table 2), we find $f_{\text{ind}} = 2\%$. The PDF for δ_0 is shown in the left panel of Fig. 6. The most probable values and 68% confidence intervals obtained are $\{\lambda_0, R_0, \delta_0, \delta, V_a\} = \{30^{+2}_{-2}, 2.2^{+0.4}_{-0.6}, -0.6^{+0.2}_{-1.3}, 0.55^{+0.04}_{-0.02}, 76^{+9}_{-11}\}$, which are consistent with values found for other models, as given in Tables 1 and 2: adding a low-energy spectral break only allows us to better adjust low-energy data (figure not shown). The best-fit parameters, for which $\chi_{\text{min}}^2 = 0.86$, are reported in Table 3. The small value of χ_{min}^2 (smaller than 1) may indicate an over-adjustment, which would disfavour the model.

It is also interesting to compel δ_0 to be positive, to check whether $\delta_0 = 0$ (equivalent to Model III), $\delta_0 = \delta$ (equivalent to Model II), or any value in-between that is preferred. We find the most probable values to

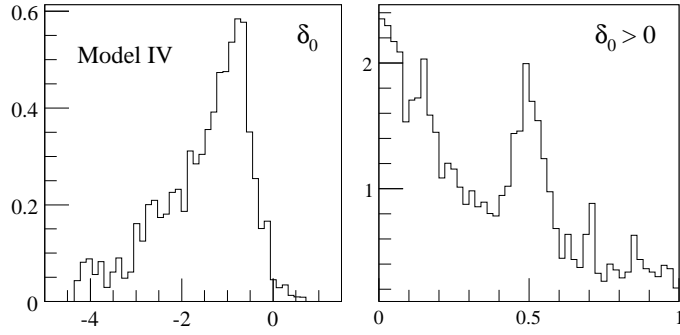


Fig. 6. Marginalised PDF for the low-energy spectral index δ_0 in Model IV-C. The parameter δ_0 is either free to span both positive and negative values (left panel) or constrained to $\delta_0 > 0$ (right panel).

be $\{\lambda_0, R_0, \delta_0, \delta, V_a\} = \{23_{-1}^{+1}, 1_{-1}^{+2}, 0^{+0.6}, 0.49_{-0.01}^{+0.01}, 102_{-5}^{+4}\}$. The corresponding PDF for δ_0 is shown in the right panel of Fig. 6. The maximum occurs for $\delta_0 = 0$, which is also found to be the best-fit value; we checked that the best-fit parameters matches those given in Table 3 for Model III-C. A secondary peak appears at $\delta_0 \approx 0.5$, such as $\delta_0 \approx \delta$ corresponding to Model II. The associated χ^2_{\min} for this configuration is worse than that obtained with $\delta_0 = 0$, in agreement with the conclusion that Model III provides a closer description of the data than Model II.

5.1.4. Summary and confidence levels for the B/C ratio

In the previous paragraphs, we have studied several models and B/C datasets. The two main conclusions that can be drawn are i) the best-fit model is Model III, which includes reacceleration and a cut-off rigidity, and ii) the most likely values of the propagation parameters are not too dependent on the data set used, although when data are inconsistent with each other the statistical interpretation of the goodness of fit of a model is altered (all best-fit parameters are gathered in Table 3). The values of the derived propagation parameters are close to the values found in similar studies and the correlation between the LB transport parameters are well understood.

Taking advantage of the knowledge of the χ^2 distribution, we can extract a list of configurations, i.e. a list of parameter sets, based on CLs of the χ^2 PDF (as explained in App. A.2). The χ^2 distribution is shown for our best model, i.e. Model III, in Fig. 7. The red and black areas correspond to the 68% and 95% confidence intervals, which are used to generate two configuration lists, from which 68% and 95% CLs on, e.g., fluxes, can be derived³.

The B/C best-fit curve (dashed blue), the 68% (red solid), and 95% (black solid) CL envelopes are shown in Fig. 8. For the specific case of the LBM, this demonstrates that current data are already able to constrain strongly the B/C flux (as reflected by the good value $\chi^2_{\min} = 1.06$), even at high energy. This provides encouraging support in the discriminating power of forthcoming data. However, this conclusion must be confirmed by analysis of a more refined

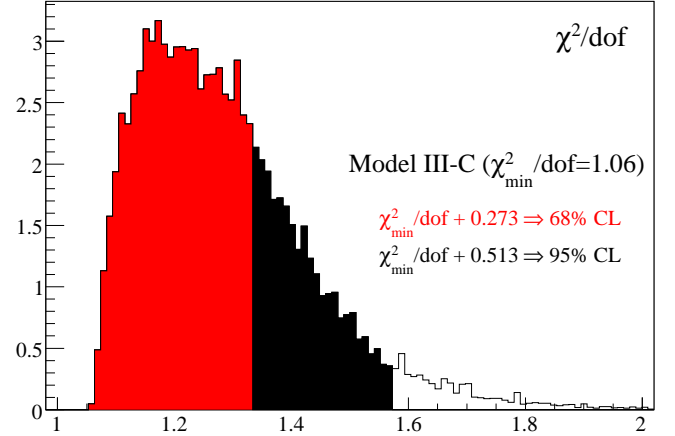


Fig. 7. χ^2/dof normalised distribution for Model III-C. The 68% and 95% CL of the distribution are shown respectively as the red and black area.

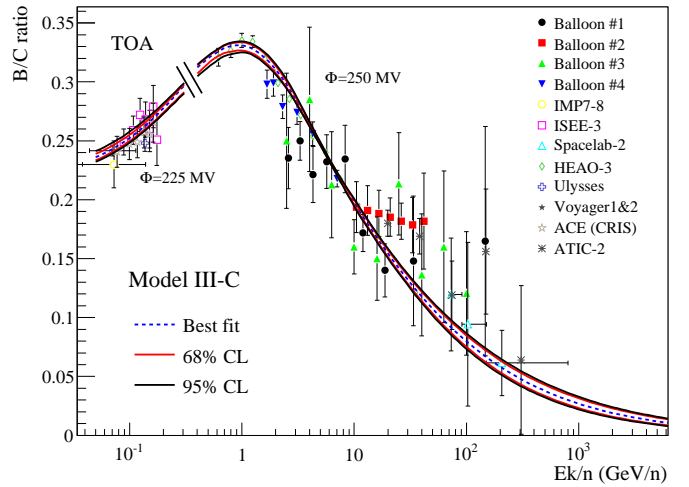


Fig. 8. Confidence regions of the B/C ratio for Model III-C as calculated from all propagation parameters satisfying Eq. (A.2). The blue-dashed line is the best-fit solution, red-solid line is 68% CL and black-solid line 95% CL. Two modulation parameters are used: $\Phi = 225$ MV below 0.4 GeV/n (adapted for ACE+Voyager 1 & 2+IMP7-8 data) and $\Phi = 250$ MV above (adapted for HEAO-3 data).

model (e.g., diffusion model), for which the situation might not be so simple.

From the same lists, we can also derive the range allowed for the source abundances of elements (we did not try to fit isotopic abundances here, although this can be achieved, e.g., as in Simpson & Connell 2001 and references therein). The element abundances are gathered in Table 4, for elements from C to Si (heavier elements were not used in this study). They can be compared with those found in Engelmann et al. (1990) (see also those derived from Ulysses data, Duvernois et al. 1996). For some elements, the agreement is striking (F, Mg), and is otherwise fair. The difference for the main progenitors of boron, i.e. C, N, and O, is a bit puzzling, and is probably related to a difference in the input-source spectral shape. This is discussed further in Sect. 5.3, where we also determine self-consistently the propagation parameters along with the C, N, and O abundances.

³ For instance, it can be used to predict the \bar{p} or \bar{d} background flux to look for a dark-matter annihilating contribution, e.g., as in Donato et al. (2004). Note however that the statistical procedure described here is far more robust than the crude approach used by Maurin et al. (2001); Donato et al. (2004).

Z	Element	$10^{22} \times q_Z$ ($\text{m}^3 \text{ s GeV/n}^{-1}$)	HEAO-3 (Engelmann et al.)
6	C	$148.5 \pm 3.$	164.9 ± 4.7
7	N	8.1 ± 0.6	9.9 ± 3.4
8	O	$185. \pm 3.$	$204. \pm 2.2$
9	F	3.67 ± 0.05	3.67 ± 0.05
10	Ne	24.1 ± 0.4	22.5 ± 1.3
11	Na	1.88 ± 0.08	1.25 ± 0.5
12	Mg	40.3 ± 0.6	40.3 ± 1.0
13	Al	3.69 ± 0.1	3.02 ± 0.6
14	Si	38.8 ± 0.5	38.8 ± 0.5

Table 4. Element source abundances, $q_Z \equiv \sum_{i=\text{isot.}} q_i$, for Model III-C (isotopic fractions are fixed to SS ones, Lodders 2003). The central values correspond, for the best-fit model, to abundances rescaled to match HEAO-3 data at 10.6 GeV/n. The uncertainty in q_i originates from the same rescaling, but arising from all combinations of parameters satisfying the 68% CL on the χ^2 distribution. For HEAO-3, the numbers are taken from Table 7 of Engelmann et al. (1990), and have been rescaled to $q_Z(\text{Si}) = 38.8 \times 10^{-22} (\text{m}^3 \text{ s GeV/n})^{-1}$ to ease the comparison.

5.2. Constraints from \bar{p}

In the context of indirect dark-matter searches, the anti-matter fluxes (\bar{p} , \bar{d} and e^+) are used to look for exotic contributions on top of the standard, secondary ones.

The standard procedure is to fit the propagation parameters to B/C data, and apply these parameters in calculating the secondary and primary (exotic) contributions. The secondary flux calculated for our best-fit Model III-C is shown, along with the data (see App. B.2 for more details) in Fig. 9 (black-solid line). For this model, we can calculate the χ^2 value for the \bar{p} data, and we find $\chi^2/\text{dof}=1.86$. The fit is not perfect, and as found in other studies (e.g., Duperray et al. 2005), the flux is somehow low at high energy (PAMELA data are awaited to confirm this trend). However, we checked that these high-energy data points are not responsible for the *large* χ^2 value. The latter could be attributed to either a small exotic contribution, a different propagation history for species for which $A/Z = 1$ or $A/Z \approx 2$, or inaccurate data.

It may therefore appear reasonable to fit directly the propagation parameters to the \bar{p} flux, assuming that it is a purely secondary species. Since the fluxes of its progenitors (p and He) are well measured, this should provide an independent check of the propagation history. We first attempted to apply the MCMC method to the \bar{p} data with Model III, then Model II and finally Model I. However, even the simplest model exhibits strong degeneracies, and the MCMC chains could not converge. We had to revert to a model with no reacceleration ($\mathcal{V}_a = 0$), no critical rigidity ($R_0 = 0$), and no break in the spectral index ($\delta_0 = 0$), for which $\lambda_{\text{esc}} = \lambda_0 \beta (R/1\text{GV})^{-\delta}$ (hereafter Model 0). The 1σ values found for the two parameters $\{\lambda_0, \delta\}$ are $\lambda_0^{\bar{p}, \text{Model 0}} = 10.2_{-0.5}^{+0.5} \text{ g} \cdot \text{cm}^{-2}$ and $\delta^{\bar{p}, \text{Model 0}} = 0.00 \pm 0.04$. Hence, only one parameter (λ_0) is required to reproduce the data, as seen in Fig. 9 (red-dashed line, $\chi^2_{\text{min}}/\text{dof}=1.128$). This is understood as follows: due to the combined effect of modulation and the tertiary contribution (\bar{p} inelastically interacting on the ISM, but surviving as a \bar{p} of lower energy), the true low-energy data points all correspond to \bar{p} produced at a few GeV energies. Due to the large scattering

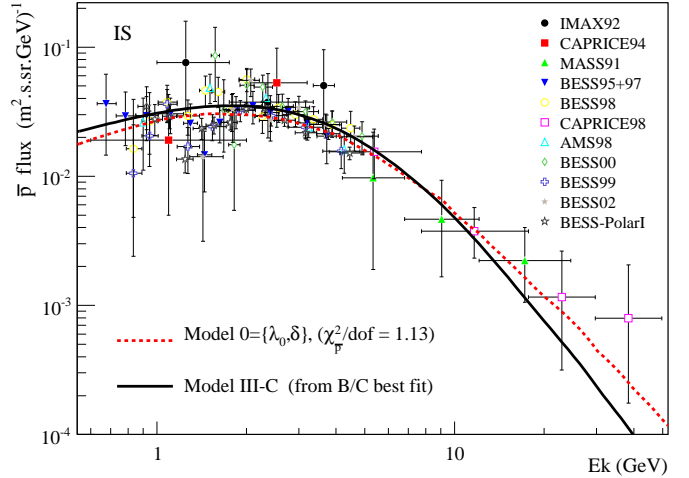


Fig. 9. Demodulated anti-proton data and IS flux for Model 0 (best fit on \bar{p} data, red-dashed line) and for Model III-C (from the best-fit parameters on B/C data, black-solid line).

in the data, it is sufficient to produce the correct amount of \bar{p} at this particular energy to account for all of the data.

Due to the importance of antimatter fluxes for indirect dark-matter searches, this novel approach could be helpful in the future. However, this would require a more robust statistics of the \bar{p} flux, especially at higher energy, to lift the degeneracy in the parameters.

5.3. Adding free parameters related to the source spectra

In all previous studies (e.g., Jones et al. 2001), the source parameters were investigated after the propagation parameters had been determined from the B/C ratio (or other secondary to primary ratio). We propose a more general approach, where we fit simultaneously all of the parameters. With the current data, this already provides strong constraints on the CR source slope α and source abundances (CNO). Higher-quality data are awaited to refine this analysis. We also show how this approach can help to uncover inconsistencies in the measured fluxes.

For all models below, taking advantage of the results obtained in Sect. 5.1, we retain Model III-C. The roman number refers to the free transport parameters of the model ($\text{III} = \{\lambda_0, R_0, \delta, \mathcal{V}_a\}$), and the capital refers to the choice of the B/C dataset ($\text{C} = \text{HEAO-3} + \text{Voyager 1 \& 2} + \text{IMP7-8}$, see Table 2). This is supplemented by source spectra parameters and additional data for the element fluxes.

5.3.1. Source shape α and η from Eq. (20)

As a free parameter, we first add a universal source slope α . We then allow η , parameterising a universal low-energy shape of all spectra, to be a second free parameter. In addition to B/C constraining the transport parameters, some primary species must be added to constrain α and η . We restrict ourselves to O, the most abundant boron progenitor, because it was measured by both the HEAO-3 experiment (Engelmann et al. 1990), and also the TRACER experiment (Ave et al. 2008). The modulation levels were $\Phi = 250$ MV for HEAO-3 and $\Phi = 500$ MV for TRACER.

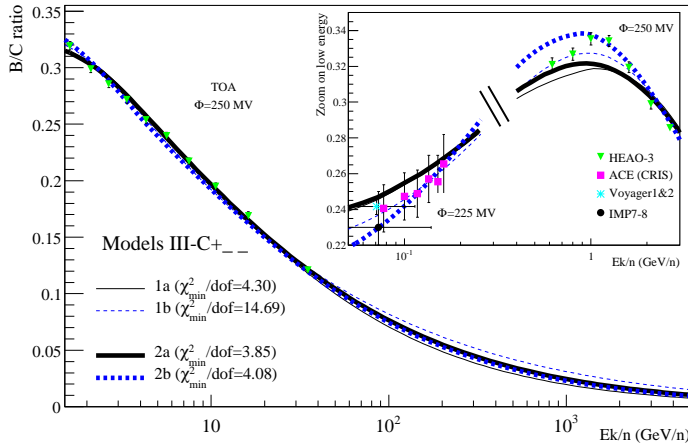


Fig. 10. B/C ratio from best-fit models of Table 6. In addition to the propagation parameters, free parameters of the source spectra are α (thin lines, labeled 1) or α and η (thick lines, labeled 2): The two models are tested on two datasets for the primary flux (in addition to using the B/C HEAO-3 data): O is as measured by HEAO-3 (black lines, dataset a) or as measured by TRACER (blue lines, dataset b).

We estimated the latter number from the solar activity at the time of flight (2 weeks in December 2003) as seen from neutron monitors data⁴.

In total, we test four models (denoted by 1a, 1b, 2a, and 2b for legibility):

- III-C+1a: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha\}$, with O=HEAO-3;
- III-C+1b: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha\}$, with O=TRACER;
- III-C+2a: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, \eta\}$, with O=HEAO-3;
- III-C+2b: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, \eta\}$, with O=TRACER;

where the Arabic numbers relate to the source-spectrum free parameters used in the calculation, and the lower case relates to the chosen oxygen-flux dataset (a=HEAO-3, b=TRACER). The most probable parameters are gathered in Table 5, where, to provide a comparison, the first line reports the values found for Model III-C (i.e. with $\gamma \equiv \alpha + \delta$ fixed to 2.65). We remark that by adding HEAO-3 oxygen data to the fit (1a), the propagation parameters λ_0 , R_0 , and δ overshoot Model III-C’s results, while they undershoot those of Model 1b (TRACER data). The parameter \mathcal{V}_a undershoots and overshoots for these two models respectively, since it is anti-correlated with the former parameters. As a consequence, the fit to B/C is worsened, especially at low energy (see Fig. 10).

The top left panel of Fig. 11 shows the slopes α derived for Models 1a (solid black) and 1b (dashed blue). In both cases, α is well constrained, but the values are inconsistent, a result that is clear because the low-energy data are also inconsistent: the demodulated (i.e. IS) HEAO-3 and TRACER oxygen data points are shown in the right panel of Fig. 11. To remedy this situation, we allow η to be a free parameter (family of models III+2). The net effect is to absorb any uncertainty originating in either the modulation level or the source-spectrum low-energy shape. As shown in the bottom panel of Fig. 11, the source slopes derived

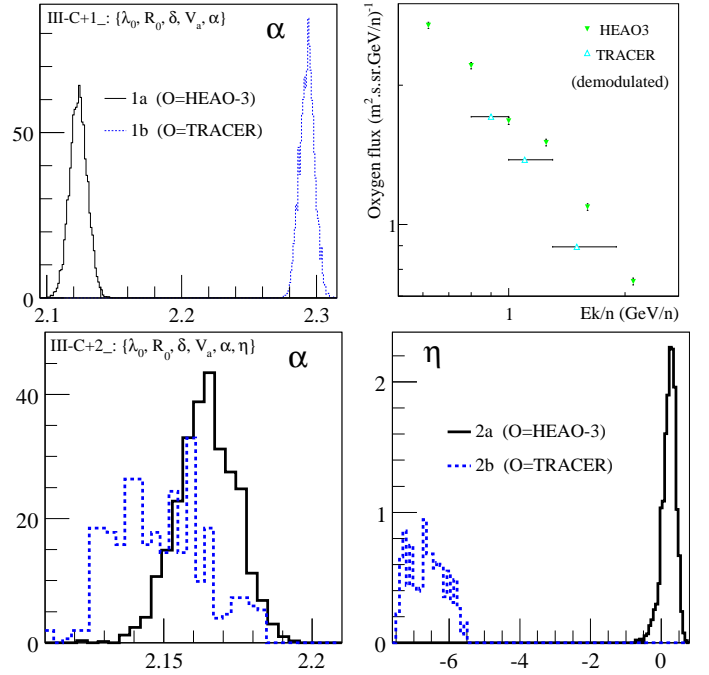


Fig. 11. Marginalised PDF for models III-C+1 (α free parameter, top left) and III-C+2 (α and η free parameters, bottom panels). In the three panels, the solid-black lines rely on HEAO-3 oxygen data, whereas the blue-dashed lines rely on TRACER oxygen data (thin and thick lines are as in Fig. 10). Top right panel: zoom in demodulated O low energy HEAO-3 and TRACER data. (The solid segments on TRACER data show the energy bin size. Uncertainty on all fluxes are present, but too small to notice).

from the two experiments are now in far closer agreement (bottom left), with $\alpha \simeq 2.15$. The most probable values and the best-fit model values are given in Tables 5 and 6 respectively. The effect of this action is evident in the low-energy slope of the source spectrum η . As seen in the bottom-right panel, the two data sets contain significantly inconsistent ranges. The value $\eta_{\text{TRACER}} \simeq -6.7$ probably indicates that the solar modulation we chose was incorrect. The value $\eta_{\text{HEAO-3}} \simeq 0.3$ might provide a reasonable guess of the low-energy shape of the source spectrum, but might also be a consequence of systematics in the experiment. The associated oxygen fluxes are shown in Fig. 12 for the best-fit models: as explained, models that allow η to vary (thick lines) reproduce more accurately the data than when η is set to be -1 (thin lines).

Although it would be precipitate to draw any firm conclusion about the low-energy shape, we can turn the argument around to serve as a diagnosis of the low-energy data quality. For instance, assuming that the spectral index α of all elements was the same, extracting and comparing η_i for each of these i elements may enable us to infer some systematics remaining in the data. It would be worth fitting the H and He species, which are the most reliably measured fluxes to date; this will be considered in a future study using diffusion models.

5.3.2. α , η and source normalisation q_i

The final two models add, as free parameters, the CNO element source abundances (relative isotopic abundances

⁴ <http://ulysses.sr.unh.edu/NeutronMonitor/Misc/neutron2.html>. Indeed, the solar activity between 2002 and 2004 has not varied much, so that we use a value for Φ derived from the BESS 2002 flight (see Fig. 2 of Shikaze et al. 2007).

Model-Data	λ_0 g cm^{-2}	R_0 GV	δ	\mathcal{V}_a $\text{km s}^{-1}\text{kpc}^{-1}$	α	η	$10^{20} \times (q_C q_N q_O)^\dagger$ $(\text{m}^3 \text{ s GeV/n})^{-1}$
III-C [‡]	27^{+2}_{-2}	$2.6^{+0.4}_{-0.7}$	$0.53^{+0.02}_{-0.03}$	86^{+9}_{-5}	-	-	-
III-C+1a	37^{+2}_{-2}	$4.4^{+0.1}_{-0.2}$	$0.61^{+0.01}_{-0.01}$	64^{+4}_{-4}	$2.124^{+0.005}_{-0.007}$	-	-
III-C+1b	$20.9^{+0.2}_{-0.8}$	$0.3^{+0.6}_{-0.1}$	$0.47^{+0.01}_{-0.01}$	103^{+2}_{-3}	$2.294^{+0.004}_{-0.006}$	-	-
III-C+2a	29^{+2}_{-2}	$2.7^{+0.3}_{-0.4}$	$0.55^{+0.01}_{-0.02}$	84^{+4}_{-7}	$2.16^{+0.01}_{-0.01}$	$0.3^{+0.1}_{-0.2}$	-
III-C+2b	32^{+4}_{-1}	$4.3^{+0.3}_{-0.1}$	$0.56^{+0.03}_{-0.01}$	62^{+2}_{-2}	$2.14^{+0.03}_{-0.01}$	$-6.7^{+0.9}_{-0.1}$	-
III-C+4a	40^{+3}_{-1}	$4.6^{+0.2}_{-0.1}$	$0.64^{+0.01}_{-0.02}$	58^{+2}_{-5}	$2.13^{+0.01}_{-0.01}$	-	$1.93^{+0.04}_{-0.004} 0.089^{+0.007}_{-0.005} 2.42^{+0.04}_{-0.05}$
III-C+5a	38^{+1}_{-2}	$4.4^{+0.1}_{-0.3}$	$0.60^{+0.02}_{-0.01}$	81^{+4}_{-1}	$2.17^{+0.02}_{-0.02}$	$-0.4^{+1.2}_{-0.1}$	$2.2^{+0.2}_{-0.1} 0.107^{+0.01}_{-0.006} 2.7^{+0.3}_{-0.1}$

[‡] III-C: propagation parameters are $\{\lambda_0, R_0, \delta, \mathcal{V}_a\}$ and the B/C dataset is HEAO-3+Voyager 1 & 2+IMP7-8.

[†] Abundances are 1.65|0.10|2.04 for HEAO-3 (Engelmann et al. 1990, and see Table 4).

Table 5. Most probable values of the propagation parameters (after marginalising over the other parameters) for models III-C+... The additional free parameters and data “...” correspond to: 1a/1b= $\{\alpha\}$, 2a/2b= $\{\alpha, \eta\}$, 4a/4b= $\{\alpha, q_C, q_N, q_O\}$, 5a/5b= $\{\alpha, \eta, q_C, q_N, q_O\}$ with either O data is HEAO-3 (a) or TRACER (b). The uncertainty on the parameters correspond to 68% CL of the marginalised PDF (see App. A). The associated best-fit parameters are gathered in Table 6.

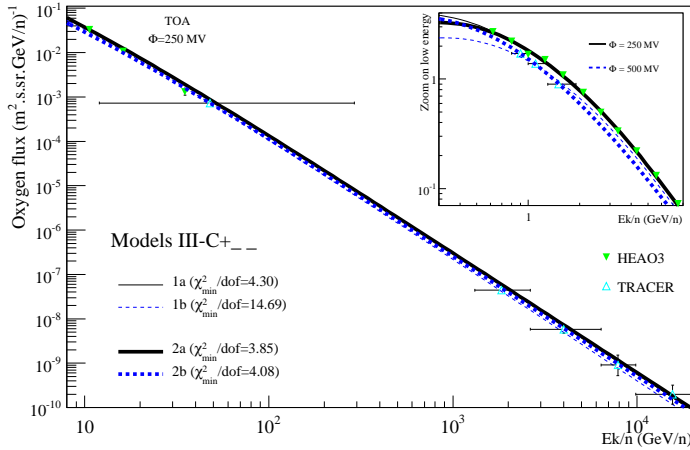


Fig. 12. Same models as in Fig. 10, but for the oxygen flux.

are fixed to SS ones). The data used in the fit are B/C, C, N, and O, all measured by HEAO-3 (TRACER data for C and N have not yet been published). The models, which are denoted by short 4a and 5a in the text below, are:

- III-C+4a: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, q_C, q_N, q_O\}$;
- III-C+5a: $\{\lambda_0, R_0, \delta, \mathcal{V}_a\} + \{\alpha, \eta, q_C, q_N, q_O\}$.

The PDF and 2D correlations plots between propagation and source parameters are seen in Fig. 13. With nine parameters, the efficiency is very low ($f_{\text{ind}} \lesssim 0.05\%$), even using the BSP trial function. To obtain ~ 800 independent points, a total $1.6 \cdot 10^6$ steps were completed. The contours are not as regular as for our 4-parameter model (see Fig. 2), but correlations between the parameters are still clearly evident: we recover the $\lambda_0 - R_0 - \delta$ correlations (and anti-correlation with \mathcal{V}_a). In addition, we note that all source-related parameters (α, η and element abundances $q_{C,N,O}$) are correlated among themselves and with \mathcal{V}_a (hence anti-correlated with the remaining transport parameters). This is especially visible for the primary species C and O, while less clear for the mixed species N. This is understood as the parameter γ , i.e. the slope of the propagated primary fluxes ($\gamma = \alpha + \delta$) is mostly fixed by measurements: if we decrease δ , then α must be increased to match the data. However,

if the source slope is increased, the exponent η must also increase to match the low-energy data. The positive correlation between source abundances comes from the fact that relative fluxes should be preserved.

The most probable values are gathered in Table 5. Compared with the respective Models 1a and 2a, leaving the source abundances q_C, q_N and q_O free in 4a and 5a does not significantly change our conclusions. Again, adding η (2a and 5a) as a free parameter allows us to absorb the low-energy uncertainties in the data, so that we obtain $\alpha = 2.17$ (5a) instead of the biased value of 2.13 (4a). The same conclusions hold for other propagation parameters. On the derived source abundances, the impact of adding the parameter η is for them to increase. The relative C:N:O abundances ($O \equiv 1$) are respectively 0.78 : 0.36 : 1 (4a) and 0.82 : 0.40 : 1 (5a), the second model providing values slightly closer to those derived from HEAO-3 data 0.81 : 0.49 : 1.

The difference in the source element abundances when they are rescaled to match the data or including them in the MCMC is also seen from Table 6, which gathers the best-fit parameters. The next-to-last line reproduces $q_{C,N,O}$ obtained for all models: all abundances are roughly in agreement, although our approach underlines the importance of taking the correlations between the parameters properly into account in extracting unbiased estimates of the propagation and source parameters.

The goodness of fit for the models when applied to the B/C, C, N, and O data is shown in the last column of Table 6, in terms of the χ^2_{min} value. The models in which $q_{C,N,O}$ is free do not provide a closer match between models and data but also a no poorer fit than when $q_{C,N,O}$ is fixed. As soon as primary fluxes are included in the fit (compared to Model III-C), the χ^2_{min} is worsened. This is due to a combination of an imperfect fit to the primary fluxes and, as already said, a poorer B/C fit because the propagation parameters are optimised to match the former rather than the latter (B/C). The best-fit parameters are given in the same Table, and the associated CNO fluxes are plotted in Fig. 14 for illustration purposes.

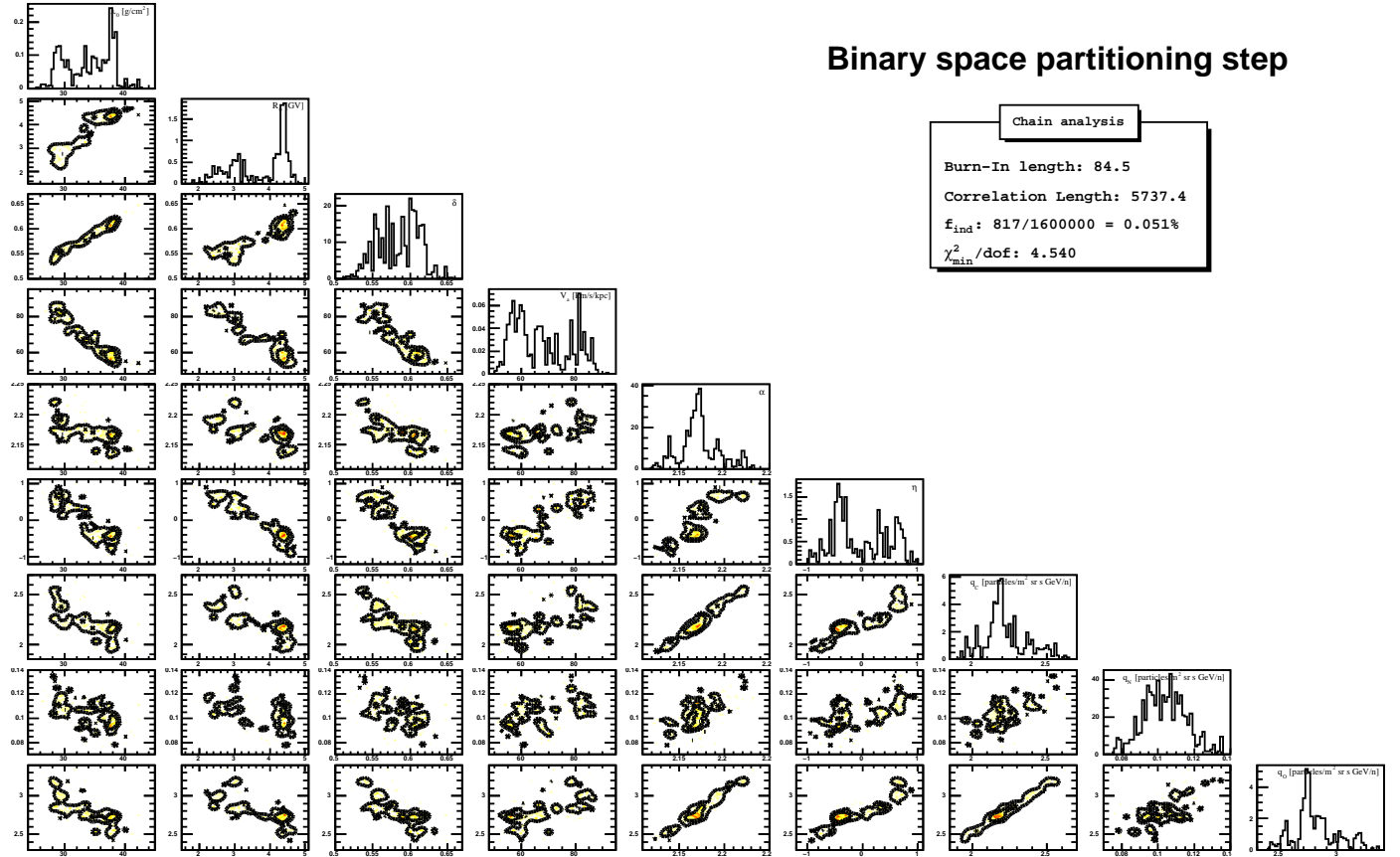


Fig. 13. PDF (diagonal) and 2D correlations (off-diagonal) plots for the nine free parameters of Model III-C+5 when fitted on B/C, and CNO HEAO-3 data.

Model-Data	λ_0^{best} g cm ⁻²	R_0^{best} GV	δ^{best}	V_a^{best} km s ⁻¹ kpc ⁻¹	α^{best}	η^{best}	$10^{20} \times (q_C q_N q_O)^{\dagger}$ (m ³ s GeV/n) ⁻¹	χ^2/dof
III-C	26.9	2.45	0.527	88.5	-	-	[1.48 0.08 1.85] [‡]	1.06
III-C+1a	36.9	4.34	0.610	64.6	2.123	-	[1.92 0.105 2.40] [‡]	4.30
III-C+1b	20.7	0.46	0.470	102.9	2.293	-	[3.43 0.219 4.12] [‡]	14.69
III-C+2a	28.7	2.61	0.547	84.5	2.168	0.305	[2.25 0.126 2.81] [‡]	3.85
III-C+2b	33.0	4.24	0.568	61.5	2.154	-6.545	[2.09 0.161 2.17] [‡]	4.08
III-C+4a	39.2	4.60	0.626	59.2	2.126	-	1.92 0.090 2.42	4.65
III-C+5a	28.6	2.44	0.545	83.0	2.175	0.449	2.27 0.104 2.86	4.54

[‡] These values are not extracted from the PDF: they are values of rescaled abundances required to match HEAO-3 CNO data at 10.6 GeV/n.

[†] For a comparison, HEAO-3 abundances (Engelmann et al. 1990, and see Table 4) are 1.65|0.10|2.04.

Table 6. Best-fit values (corresponding to χ^2_{min}) for all models as given in Table 5. Number of data points for the $\chi^2_{\text{min}}/\text{dof}$ calculation: 22 B/C data (III-C=HEAO-3+Voyager 1 & 2+IMP7-8) plus 14 oxygen HEAO-3 data for 1a and 2a, 8 oxygen TRACER data for 1b and 2b, or 14×3 (C, N and O) HEAO-3 data for 4a and 5a.

5.3.3. Perspective on source spectrum parameters

Other primary species could have been included in the χ^2 calculation to i) constrain further α , and/or ii) to check the hypothesis $\alpha_i \neq \alpha_j$ for different species, and/or iii) diagnose some problems in the data, if we believe the slope should be universal. However, using a few primary species (O or CNO) already affects the goodness of fit of B/C (compare model III-C to others in Fig. 10). Since there are many more measured primary fluxes than secondary ones, taking too many primaries would weigh too significantly in the

χ^2 , compared to the B/C contribution, and this would divert the MCMC into regions of the parameter space that fit these fluxes rather than B/C. Since systematic errors are known to be larger in flux measurements than in ratios, this may lead to biased estimates of the propagation parameters. Allowing η to be a free parameter is a first step to decrease the impact of this bias (in the low-energy part of the spectrum). These biases are not necessarily an issue since we may be more interested in estimates of the source parameters rather than in unbiased value of the propagation parameters.

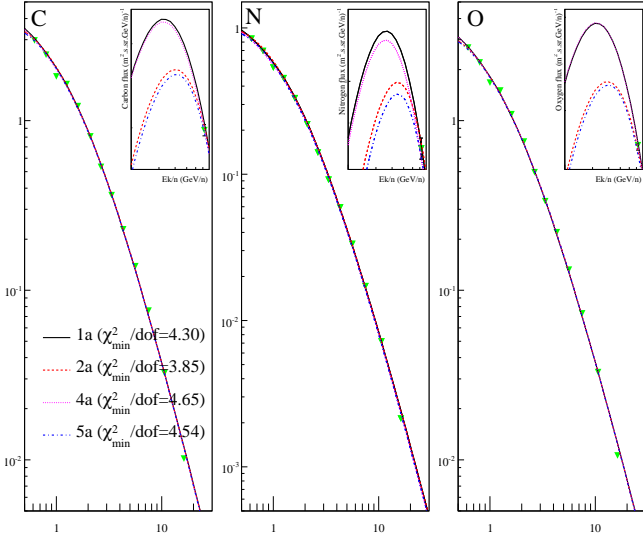


Fig. 14. Carbon, nitrogen and oxygen fluxes from best-fit models of Table 6.

To illustrate the difficulty in data systematics, it suffices to say that for the most abundant species H and He, all experiments provided mutually incompatible measurements, until AMS and BESS experiments flew ten years ago. We cannot therefore expect HEAO-3 data, acquired in 1979, to be completely free of such drawbacks. Again, we await the publication of several forthcoming new data sets before pursuing our analysis further in this direction.

6. Conclusion

We have implemented a Markov Chain Monte Carlo to extract the posterior distribution functions of the propagation parameters in a LBM. Three trial functions were used, namely a standard Gaussian step, an N-dimensional Gaussian step and its covariance matrix, and a binary-space partitioning. For each method, a large number of chains were processed in parallel to accelerate the PDF calculations. The three trial functions were used sequentially, each method providing some inputs to the next: while the first one was good at identifying the general range of the propagation parameters, it was not as efficient in providing an accurate description of the PDF. The two other methods provided this accuracy, and the final results were based on PDF obtained from the chains processed with the binary-space partitioning.

Taking advantage of the sound statistical properties of the MCMC, confidence intervals for the propagation parameters can be given, as well as confidence contours for all fluxes and other quantities derived from the propagation parameters. The MCMC was also used to compare the impact of choosing different datasets and ascertain the merits of different hypotheses concerning the propagation models. Concerning the first aspect, we have shown that combining different B/C datasets leaves mostly unchanged the propagation parameters, while strongly affecting the assessment of the goodness of a model. We also show that at present, the \bar{p} data do not cover a sufficiently large energy range to constrain the propagation parameters, but they could be useful in crosschecking the secondary nature of the flux in the future.

In this first paper, we have focused on the phenomenologically well-understood LBM, to ease and simplify the discussion and implementation of the MCMC. In agreement with previous studies, we confirm that a model with a rigidity cutoff performs more successfully than one without and that reacceleration is preferred over no reacceleration. Such a model can be associated with a diffusion model with wind and reacceleration. As found in Maurin et al. (2001), the best-fit models demand both a rigidity cutoff (wind) and reacceleration, but do not allow us to reconcile the diffusion slope with a Kolmogorov spectrum for turbulence. An alternative model with two slopes for the diffusion was used, but it is not favoured by the data. In a last stage, we allowed the abundance and slope of the source spectra to be free parameters, as well as the element abundances of C, N, and O. This illustrated a correlation between the propagation and source parameters, potentially biasing the estimates of these parameters. The best-fit model slope for the source abundances was $\alpha \approx 2.17$ using HEAO-3 data, compatible with the value $\alpha \approx 2.14$ for TRACER data. The MCMC approach allowed us to draw confidence intervals for the propagation parameters, the source parameters, and also for all fluxes.

A wealth of new data on Galactic CR fluxes are expected soon. As illustrated for the LBM, the MCMC is a robust tool in handling complex data and model parameters, where one has to fit simultaneously all source and propagation parameters. The next step is to apply this approach to more realistic diffusion models and larger datasets, on a wider range of nuclear species.

Acknowledgements. D.M. and R.T. warmly thank Joanna Dunkley for useful discussions at an early stage of this work. We thank Céline Combet for a careful reading of the paper.

Appendix A: Best fit, goodness of a model, most probable values and confidence levels/intervals

The best-fit model parameters are given by a unique set of parameters for which the χ^2 value of Eq. (13) is minimized; the goodness of fit of a model is given by χ^2_{\min}/dof . On the other hand, the most probable value for each parameter θ_i is defined as the maximum $\mathcal{P}_i^{\max} \equiv \mathcal{P}(\theta_i^{\max})$ of its PDF (after marginalising). The most probable θ_i^{\max} and best-fit model parameters θ^{best} do not necessarily coincide, especially when correlations exist between parameters. The best-fit model parameters are best suited to providing the most likely CR fluxes, whereas the 1D marginalised PDF provides directly the most likely value of the parameter.

A.1. Confidence levels/intervals on parameters

Confidence intervals (CI), associated with a confidence level (CL), are constructed from the PDF. The asymmetric interval $\Delta_x \equiv [\theta_i^{\max} - \theta_x^-, \theta_i^{\max} + \theta_x^+]$ such as

$$\text{CL}(x) \equiv \int_{\Delta_x} \mathcal{P}(\theta_i) d\theta_i = 1 - \gamma, \quad (\text{A.1})$$

defines the $1 - \gamma$ confidence level (CL), along with the CI of the parameter θ_i . Here, the CIs (i.e. θ_x^- and θ_x^+) are found by decreasing the value $\mathcal{P}(\theta_i)$ from \mathcal{P}_i^{\max} to \mathcal{P}_i^x , such that $1 - \gamma = x$. This is easily generalised to 2D confidence levels in constructing 2D confidence intervals as shown later in correlation plots. Below, we use the $x = 68\%$ and $x = 95\%$ CLs, corresponding to 1σ and 2σ uncertainties.

A.2. Confidence intervals on fluxes

The best-fit model fluxes (e.g., B/C, O, \bar{p}) are calculated from the best-fit model parameters. Confidence levels in these quantities cannot be obtained from the 1D marginalised CIs of the parameters. They must be constructed from a sampling of the (still) correlated parameters. This is achieved by using all sets of parameters $\{\theta\}_{x\%CL} = \{\theta_i\}_{i=1\dots p}$, for which $\chi^2(\theta_i)$ falls in the $x\%$ confidence level of the χ^2 PDF. Once these sets are found, we simply calculate the desired flux for all the sets: the maximum and minimum values are kept for each energy bin, defining confidence envelopes for this flux. Thus, the main task is to construct confidence intervals for the χ^2 distribution.

For n parameters in the large sample limit—where the joint PDF for the estimator of the parameters and the likelihood function become Gaussian—the CI is given by

$$[\chi_{\min}^2, \chi_{\min}^2 + \Delta\chi^2], \quad \text{where} \quad \Delta\chi^2 = Q_\gamma(1 - \gamma, n)$$

is the quantile of order $1 - \gamma$ (confidence level CL) of the χ^2 distribution (Cowan 1997). However, by applying the MCMC, we have access to a direct sampling of the χ^2 distribution. Hence, independently of the statistical meaning of a model, the confidence interval is extracted from the cumulative χ^2 PDF, by requiring that

$$\int_{\chi_{\min}^2}^{\chi_{\min}^2 + \Delta\chi^2} \mathcal{P}(\chi^2) d\chi^2 = 1 - \gamma. \quad (\text{A.2})$$

We nevertheless checked that both approaches provide very similar results. For instance, the CIs (for Model III-C) obtained directly from Fig. 7 are CI (68%) = $[\chi_{\min}^2, \chi_{\min}^2 + 4.9]$ and CI (95%) = $[\chi_{\min}^2, \chi_{\min}^2 + 9.2]$, whereas they are CI (68%) = $[\chi_{\min}^2, \chi_{\min}^2 + 4.7]$ and CI (95%) = $[\chi_{\min}^2, \chi_{\min}^2 + 9.5]$ when calculated from the $Q_\gamma(1 - \gamma, n)$ quantiles (Cowan 1997).

Appendix B: Data

In the paper, we focus on the B/C ratio, which is the most accurate tracer of the propagation parameters (other tracers, such as the sub-Fe/Fe or the quartet ^1H , ^2H , ^3He and ^4He are not considered). We also estimate the potential of the \bar{p} , a secondary species, as an alternative species for constraining these parameters. We describe below the typical configurations used to calculate the corresponding spectrum as well as the associated datasets used.

B.1. B/C

The default configuration for nuclei is the following: the value of the observed propagated slope $\gamma = \alpha + \delta$, unless stated otherwise, is set to be 2.65 (Ave et al. 2008), and source abundances of the most abundant species (C, N, O, F, Ne, Na, Mg, Al, Si) are rescaled to match HEAO-3 data at 10.6 GeV/n. Boron is assumed to be a pure secondary species. Only elements lighter than Si are propagated, since they are the only relevant ones for determining B/C (Maurin et al. 2001).

For B/C at intermediate GeV energies, we use HEAO-3 data (Engelmann et al. 1990). They are complemented at low energy by the ACE (CRIS) data (de Nolfo et al.

2006). For a few model iterations, we also look for combined constraints from multiple sets of data. A collection of low-energy data is formed by data sets for the IMP7-8 (Garcia-Munoz et al. 1987), ISEE-3 (Krombel & Wiedenbeck 1988), Ulysses (Duvernois et al. 1996), and Voyager 1&2 (Lukasik et al. 1999) spacecrafts. At higher energy, we consider several balloon flights (Lezniak & Webber 1978; Orth et al. 1978; Simon et al. 1980; Dwyer & Meyer 1987), the ATIC-2 balloon-borne experiment (Panov et al. 2007), and the Spacelab-2 experiment (Mueller et al. 1991). For element fluxes, HEAO-3 and TRACER results (Ave et al. 2008) are used.

B.2. \bar{p}

For the calculation of the \bar{p} flux, the situation is simpler: the production of this secondary flux can be directly linked to the accurate measurements of propagated p and He fluxes, by the AMS (Alcaraz et al. 2000b,a; AMS Collaboration et al. 2000) and BESS experiments (Sanuki et al. 2000; Shikaze et al. 2007). For more details of the \bar{p} flux calculation (cross sections, source terms, ...), the reader is referred to Donato et al. (2001).

For the \bar{p} data, we consider the AMS 98 (AMS Collaboration et al. 2002) experiment on the shuttle, the balloon-borne experiments IMAX 92 (Mitchell et al. 1996), CAPRICE 94 (Boezio et al. 1997), WIZARD-MASS 91 (Basini 1999), CAPRICE 98 (Boezio et al. 2001), and the series of BESS balloon flights BESS 95+97 (Orito et al. 2000), BESS 98 (Maeno et al. 2001), BESS 99 and 2000 (Asaoka et al. 2002), and BESS 2002 (Haino & et al. 2005). We also add the BESS Polar results (BESS Collaboration et al. 2008). For BESS data, we use the solar modulation level as provided in Shikaze et al. (2007), based on proton and helium data.

Appendix C: Illustration of MCMC chains and PDF found with the three trial functions $q(\theta_{\text{trial}}, \theta_i)$

To compare the three trial functions (Sect. 3.3.2), a simple setup is retained: the B/C ratio observed from HEAO-3 data is used to constrain the model parameters $\{\theta_i\}_{i=1,\dots,3} = \{\lambda_0, R_0, \delta\}$, i.e. Model I.

Taking advantage of parallel processing (as underlined in Sect. 4.1), we combine several chains of 10 000 steps for each trial function. We start with the Gaussian trial function. It combines $N_c = 40$ chains and its output is used to calculate the covariance matrix. Taking advantage of smaller burn-in and correlation lengths, only 20 chains need to be combined for the covariance trial function. Again, the resulting PDFs are used as input to the BSP trial function, which also combines 20 chains. Several chains (shown here for δ), along with their burn-in length and correlation step l , are shown in Fig. C.1.

The result of the three sampling methods (Gaussian, covariance matrix, and BSP) for the PDF are shown in Fig. C.2. The insert in each panel provides mean values (over the number of chains processed N_c) for relevant parameters of the chain analysis (Sect. 3.2): a decrease in the burn-in length b (421.5, 20.4 and 2.6) and the correlation length l (159.7, 6 and 1) is found when moving from the Gaussian to the BSP sampling method. The fraction of independent samples [as defined in Eqs. (9) and (10)]

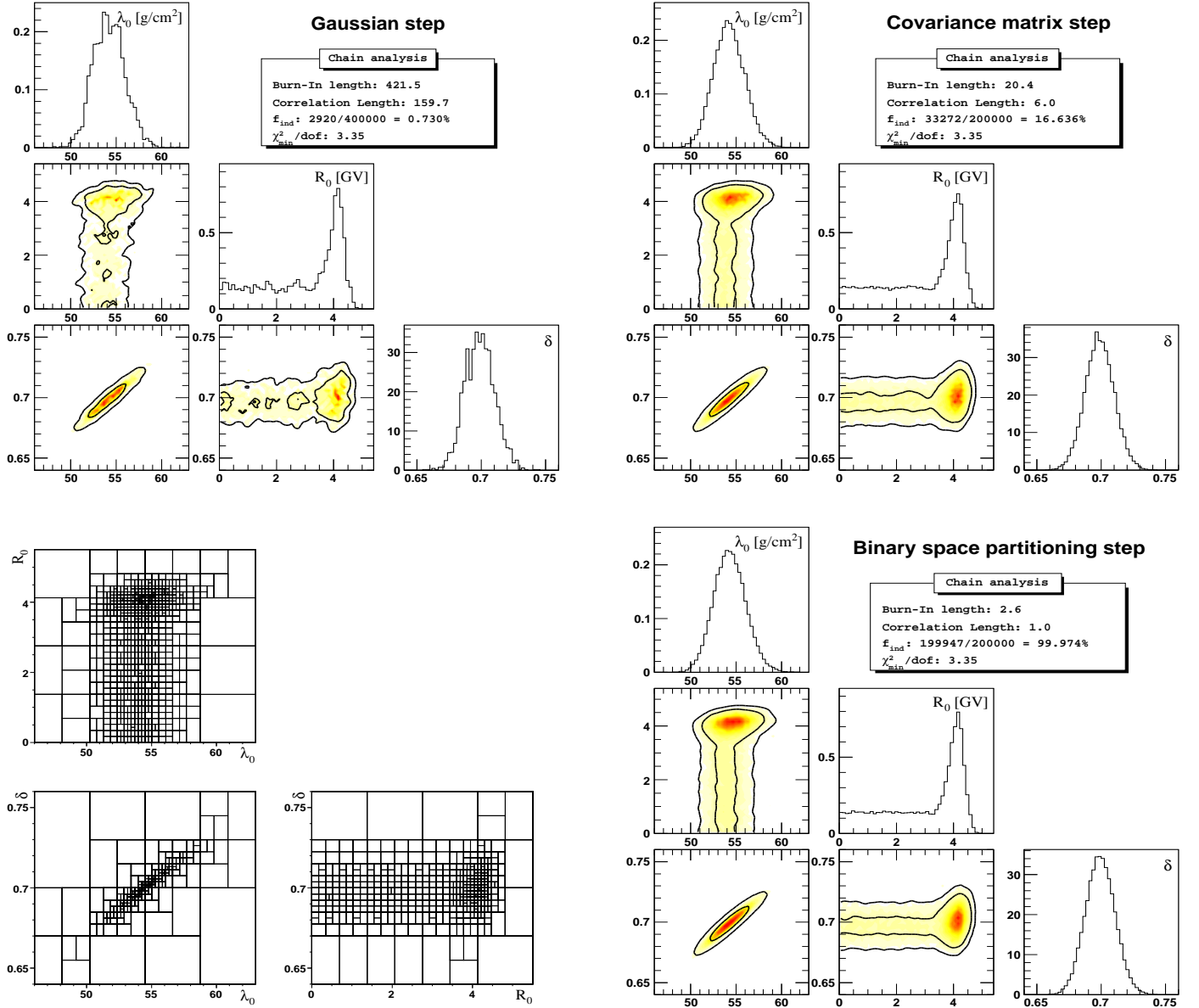


Fig. C.2. Posterior PDF of $\{\theta^{(\alpha)}\}_{\alpha=1,\dots,3} = \{\lambda_0, R_0, \delta\}$ using the 3 proposal densities defined in Sect. 3.3.2 (Gaussian—top left; Covariance matrix—top right; BSP—bottom right). In these three panels, the diagonal shows the 1D marginalised posterior density function of the indicated parameter. The number of entries in the normalised histograms corresponds to the number of uncorrelated steps spent by the chain for values of λ_0 , R_0 and δ . Off-diagonal plots show the 2D marginalised posterior density functions for the parameters in the same column and same line respectively: $\lambda_0 - R_0$, $\lambda_0 - \delta$ and $R_0 - \delta$. The colour code corresponds to the regions of increasing probability (from paler to darker shade). The two contours (smoothed) delimit regions containing respectively 68% and 95% (inner and outer contour) of the PDF. Finally, the bottom-left panel shows an example of a 3D parameter mesh (3 slices along the three different planes $\lambda_0 - \delta - R_0$) obtained from the BSP method.

is $f_{\text{ind}} = 0.7\%$ for the Gaussian step, while nearly every step is valid and uncorrelated (total of 99.9%) for the BSP mode. This confirms that, for a given number of steps, refined trial functions are more efficient in extracting the PDF (note however that some improvement comes from the fact that each method takes advantage of the previous step of calculation).

Figure C.2 (bottom-left) illustrates the binary-space partitioning discussed in Sect. 3.3.2. It shows the projections of box sides on the three 2D planes $\lambda_0 - R_0$, $\lambda_0 - \delta$ and $R_0 - \delta$ of the parameter space. The partitioning has been produced by the BSP method using the covariance matrix procedure presented on the same figure (top-right),

where the box density is clearly proportional to the estimated target density.

References

- Alcaraz, J., Alpat, B., Ambrosi, G., et al. 2000a, Physics Letters B, 490, 27
- Alcaraz, J., Alvisi, D., Alpat, B., et al. 2000b, Physics Letters B, 472, 215
- AMS Collaboration, Aguilar, M., Alcaraz, J., et al. 2002, Phys. Rep., 366, 331
- AMS Collaboration, Alcaraz, J., Alpat, B., et al. 2000, Physics Letters B, 494, 193

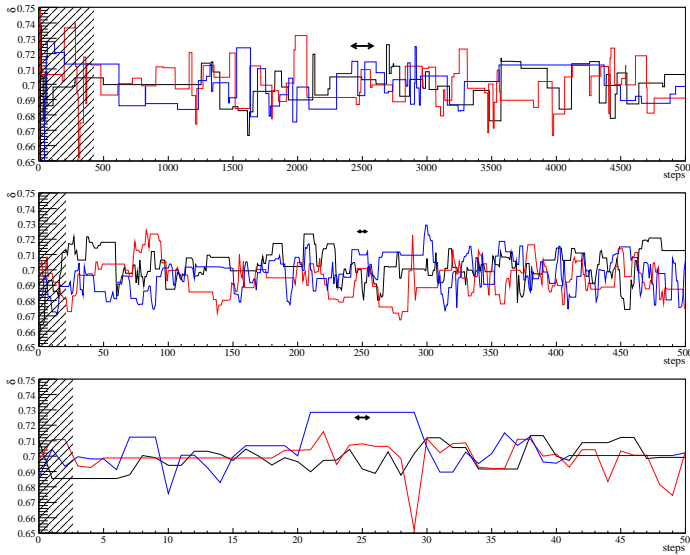


Fig. C.1. Illustration of MCMC chains (here for the parameter δ and Model I). From top to bottom, chains generated from the Gaussian step, covariance matrix, and binary-space partitioning. Three chains are shown in each panel: the shaded area corresponds to the burn-in length and the arrow to the size of the correlation length l defined by Eq. (8). Although each process consists of 10 000 steps, the Gaussian step zoom in on the first 5 000 steps, the two others displaying respectively 500 and 50 steps. This indicates that each method allows a gain of ~ 10 in efficiency to extract the PDF (compare the size of the arrows with the number of steps in each case).

Asaoka, Y., Shikaze, Y., Abe, K., et al. 2002, *Physical Review Letters*, 88, 051101
Ave, M., Boyle, P. J., Gahbauer, F., et al. 2008, *ApJ*, 678, 262
Basini, G. 1999, 3, 77
Berezko, E. G., Ksenofontov, L. T., Ptuskin, V. S., Zirakashvili, V. N., & Völk, H. J. 2003, *A&A*, 410, 189
Berezinskii, V. S., Bulanov, S. V., Dogiel, V. A., & Ptuskin, V. S. 1990, *Astrophysics of cosmic rays* (Amsterdam: North-Holland, 1990, edited by Ginzburg, V.L.)
BESS Collaboration, Abe, K., Fuke, H., et al. 2008, *Physics Letters B*, 670, 103
Binns, W. R., Garrard, T. L., Gibner, P. S., et al. 1989, *ApJ*, 346, 997
Blasi, P. 2008, *ArXiv e-prints*, 801
Bloemen, J. B. G. M., Dogiel, V. A., Dorman, V. L., & Ptuskin, V. S. 1993, *A&A*, 267, 372
Boezio, M., Bonvicini, V., Schiavon, P., et al. 2001, *ApJ*, 561, 787
Boezio, M., Carlson, P., Francke, T., et al. 1997, *ApJ*, 487, 415
Boyle, P. J., Ave, M., Gahbauer, F., et al. 2007, *astro-ph/0703707* (36th COSPAR Scientific Assembly)
Casolino, M., Picozza, P., Altamura, F., et al. 2008, *Advances in Space Research*, 42, 455
Cassé, M. & Goret, P. 1973, 1, 584
Christensen, N., Meyer, R., Knox, L., & Luey, B. 2001, *Classical and Quantum Gravity*, 18, 2677
Cowan, G. 1997, *Statistical data analysis* (Statistical data analysis. Publisher: Oxford: Clarendon Press, 1998 Series: Oxford science publications. ISBN: 0198501560)
de Berg, M., van Kreveld, M., Overmars, M., & Schwarzkop, O. 2000, *Computational Geometry* (Sec. 12: Binary Space Partitions: pp.251-265) (Springer-Verlag (2nd revised edition). ISBN : 3-540-65620-0)
de Nolfo, G. A., Moskalenko, I. V., Binns, W. R., et al. 2006, *Advances in Space Research*, 38, 1558
Delahaye, T., Lineros, R., Donato, F., Fornengo, N., & Salati, P. 2008, *Phys. Rev. D*, 77, 063527
Donato, F., Fornengo, N., Maurin, D., Salati, P., & Taillet, R. 2004, *Phys. Rev. D*, 69, 063501
Donato, F., Maurin, D., Salati, P., et al. 2001, *ApJ*, 563, 172
Dunkley, J., Bucher, M., Ferreira, P. G., Moodley, K., & Skordis, C.

2005, *MNRAS*, 356, 925
Duperray, R., Baret, B., Maurin, D., et al. 2005, *Phys. Rev. D*, 71, 083013
Duvernois, M. A., Simpson, J. A., & Thayer, M. R. 1996, *A&A*, 316, 555
Dwyer, R. & Meyer, P. 1987, *ApJ*, 322, 981
Ellison, D. C., Drury, L. O., & Meyer, J.-P. 1997, *ApJ*, 487, 197
Engelmann, J. J., Ferrando, P., Soutoul, A., Goret, P., & Juliusson, E. 1990, *A&A*, 233, 96
Evoli, C., Gaggero, D., Grasso, D., & Maccione, L. 2008, *Journal of Cosmology and Astro-Particle Physics*, 10, 18
Garcia-Munoz, M., Simpson, J. A., Guzik, T. G., Wefel, J. P., & Margolis, S. H. 1987, *ApJSupp. Series*, 64, 269
Haino, S. & et al. 2005, 3, 13
Jones, F. C. 1979, *ApJ*, 229, 747
Jones, F. C. 1994, *ApJS*, 90, 561
Jones, F. C., Lukasiak, A., Ptuskin, V., & Webber, W. 2001, *ApJ*, 547, 264
Krombel, K. E. & Wiedenbeck, M. E. 1988, *ApJ*, 328, 940
Lewis, A. & Bridle, S. 2002, *Phys. Rev. D*, 66, 103511
Lezniak, J. A. & Webber, W. R. 1978, *ApJ*, 223, 676
Lionetto, A. M., Morselli, A., & Zdravkovic, V. 2005, *Journal of Cosmology and Astro-Particle Physics*, 9, 10
Lodders, K. 2003, *ApJ*, 591, 1220
Lukasiak, A., McDonald, F. B., & Webber, W. R. 1999, 3, 41
MacKay, D. 2003, *Information Theory, Inference, and Learning Algorithms* (Publisher: Cambridge University Press. ISBN: 0521642981)
Maeno, T., Orito, S., Matsunaga, H., et al. 2001, *Astroparticle Physics*, 16, 121
Marrocchesi, P. S. et al. 2006, 36, 3129
Maurin, D., Donato, F., Taillet, R., & Salati, P. 2001, *ApJ*, 555, 585
Maurin, D., Taillet, R., & Donato, F. 2002, *Astronomy and Astrophys.*, 394, 1039
Meyer, J.-P., Drury, L. O., & Ellison, D. C. 1997, *ApJ*, 487, 182
Mitchell, J. W., Barbier, L. M., Christian, E. R., et al. 1996, *Physical Review Letters*, 76, 3057
Mueller, D., Swordy, S. P., Meyer, P., L'Heureux, J., & Grunsfeld, J. M. 1991, *ApJ*, 374, 356
Neal, R. M. 1993, *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Technical Report CRG-TR-93-1, Department of Computer Science, University of Toronto
Orito, S., Maeno, T., Matsunaga, H., et al. 2000, *Physical Review Letters*, 84, 1078
Orth, C. D., Buffington, A., Smoot, G. F., & Mast, T. S. 1978, *ApJ*, 226, 1147
Osborne, J. L. & Ptuskin, V. S. 1988, *Soviet Astronomy Letters*, 14, 132
Panov, A. D., Adams, J. H., Ahn, H. S., et al. 2006, *astro-ph/0612377*, 29-th Russian Cosmic Rays Conference
Panov, A. D., Sokolskaya, N. V., Adams, J. H., et al. 2007, *arXiv:0707.4415*, 30-th International Cosmic Rays Conference
Picozza, P., Galper, A. M., Castellini, G., et al. 2007, *Astroparticle Physics*, 27, 296
Sanuki, T., Motoki, M., Matsumoto, H., et al. 2000, *ApJ*, 545, 1135
Seo, E. S., Ahn, H. S., Beatty, J. J., et al. 2004, *Advances in Space Research*, 33, 1777
Seo, E. S. & Ptuskin, V. S. 1994, *ApJ*, 431, 705
Seo, E. S. et al. 2006, 36, 1846
Shibata, T., Hareyama, M., Nakazawa, M., & Saito, C. 2006, *ApJ*, 642, 882
Shikaze, Y., Haino, S., Abe, K., et al. 2007, *Astroparticle Physics*, 28, 154
Simon, M., Spiegelhauer, H., Schmidt, W. K. H., et al. 1980, *ApJ*, 239, 712
Simpson, J. A. & Connell, J. J. 2001, *Space Science Reviews*, 97, 337
Strong, A. W. & Moskalenko, I. V. 1998, *ApJ*, 509, 212
Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, *Annual Review of Nuclear and Particle Science*, 57, 285
Wakely, S. P. et al. 2006, 36, 3231
Webber, W. R., Kish, J. C., Rockstroh, J. M., et al. 1998, *ApJ*, 508, 940
Webber, W. R., Lee, M. A., & Gupta, M. 1992, *ApJ*, 390, 96